



UNIVERSITÀ DEGLI STUDI DI TORINO

DIPARTIMENTO DI PSICOLOGIA

DOTTORATO DI RICERCA IN NEUROSCIENZE

CICLO XXXIII

**Tracking Fetal Behavioral Development:
quantitative approaches for studying fetal movement and visual preference
in utero**

TESI PRESENTATA DA: Novella Pretti

TUTOR: Prof. Cristina Becchio

COORDINATORE DEL DOTTORATO: Prof. Andrea Calvo

ANNI ACCADEMICI: 2017-2021

SETTORE SCIENTIFICO-DISCIPLINARE DI AFFERENZA: M-PSI/01

Acknowledgments

This PhD thesis would not have been possible without the guidance and the help of my supervisor, Cristina Becchio. Thanks for providing me with the rigorous training and opportunities to make this work possible.

I would like to thank Prof. Dario Paladini, head of Fetal Medicine & Surgery Unit at Gaslini Institute, who gave me the unexpected opportunity to target my research to the fascinating and challenging field of Prenatal Medicine.

I am grateful to members of the EDL@IIT for having realized the TCFVS device (“the spaceship”).

Thank you to Jacopo Bonato for his crucial contribution in Chapter 3 and Andrea Cavallo for his (always) valuable help in Chapter 4.

I am lucky to have an incredible team of interdisciplinary researchers who made this research possible!

I would also like to express my deepest gratitude to my lab mates, Matilde, Giulia, Eugenio, Giacomo, and Enrico, and the entire C'MoN Unit. Sharing ideas and lunches with you made the lab a better place!

I express my gratitude to all the participants who attended the Gaslini Institute and took part in my studies.

Finally, I would like to thank my friends and family for their love and encouragement - especially Mum, Dad, and Edoardo - and Luca, for his full, unreserved and sharp support.

Table of Contents

List of figures	6
List of tables	7
1. Chapter 1: Introduction	8
1.1 The origin of fetal behavior	8
1.2 What we currently know about fetal behavior (and brief remarks on the prenatal neuromotor development).....	9
1.3 Why study fetal behavioral?	13
1.4 Reconceptualize fetal behavioral assessment.....	14
1.5 Towards a fetal psychology and beyond.....	18
1.6 Difficulty of measuring fetal behavior.....	22
1.7 Thesis overview	23
2. Chapter 2: Why 4D ultrasound has not (yet) revolutionized fetal movement research	26
2.1 Challenges involved in 4D US assessment of fetal behavior	27
2.2 Systematic review of 4D US fetal movement studies	30
2.2.1 Assessing the frequency of fetal movements	31
2.2.2 Assessing the quality of fetal movements	32
2.2.3 Assessing the kinematics of fetal movements.....	33
2.3 Future is now.....	34
2.4 Supporting Information.....	35
2.4.1 Methods.....	35
2.4.1.1 Protocol and registration	35
2.4.1.2 Eligibility criteria	36
2.4.1.3 Information sources and data extraction	37
2.4.2 Results.....	38
2.4.2.1 Literature search and study characteristics.....	38

2.4.2.2 Narrative synthesis of studies reporting on the frequency of fetal movements	41
2.4.2.3 Narrative synthesis of studies assessing the quality of fetal movements	43
2.4.2.4 Narrative synthesis of studies assessing the kinematics of fetal movements	44
3. Chapter 3: Towards the Use of 2D Video-Based Markerless Motion Capture to Track Fetal Movements.....	51
3.1 Introduction.....	51
3.2 Methods.....	53
3.2.1 Data resource	53
3.2.1.1 Study population	53
3.2.1.2 Image data collection	54
3.2.1.3 Image data annotation	54
3.2.1.4 Generation of a <i>consensus annotation</i>	55
3.2.1.5 Measuring performance against <i>consensus annotation</i>	56
3.2.1.6 Algorithm and human performances evaluation against <i>consensus annotation</i>	57
3.3 Results.....	57
3.3.1 Comparing DeepLabCut against manual annotations.....	57
3.4 Comparison to a fourth human labeler.....	62
3.5 Use case: extraction of wrist velocity	64
3.6 Discussion	66
3.7 Conclusion	69
3.8 Author contribution.....	70
4. Chapter 4: Fetal Visual Perception and Biological Motion Preference	71
4.1 Introduction.....	71
4.1.1 Prenatal visual development	71
4.1.2 Why biological motion	73
4.1.3 Our pilot study	74

4.2 Methods.....	75
4.2.1 Participants.....	75
4.2.2 Data acquisition	76
4.2.3 Apparatus and stimuli	76
4.2.3.1 Light source	76
4.2.3.2 Light stimulus configuration	78
4.2.4 Experimental design.....	80
4.2.5 Procedure	81
4.2.6 Data coding	82
4.2.7 Data analysis	83
4.3 Results.....	84
4.4 Discussion	87
4.5 Conclusion	90
5. Chapter 5: Discussion	92
5.1 Summary of the main findings.....	92
5.2 Clinical considerations.....	95
5.3 Outlook	97
6. References.....	100

List of Figures

Figure 1.1 Emergence of specific fetal movement patterns with time presented as median and range of the first appearance during serial observations on individuals. Reproduced from (de Vries et al., 1982).	12
Figure 2.1 Impact of fetal growth on visualizing fetal movements at a fixed ROI volume.....	29
Figure 2.2 Impact of frame rate on visualizing fetal movements.....	30
Figure S1 PRISMA flowchart.....	40
Figure 3.1 Examples of frames and labeled bodyparts predicted by DeepLabCut	59
Figure 3.2 Mean F1 scores of predicted and human labelers' detections of each key point computed against the consensus annotations	60
Figure 3.3 Averaged root mean square error (RMSE) of predicted and human labelers' positions of each key point computed against the consensus annotations	60
Figure 3.4 Example of trajectory and velocity profiles based on DeepLabCut predictions.....	65
Figure 4.1 Coronal section of the fetal head at the orbital level in a 30-week-old fetus where the lens of the eyes (white circles) are easily visualized	75
Figure 4.2 A) LEDs matrix display generating visual stimuli. B) Biological movement profile recording with an optical motion capture system.....	78
Figure 4.3 Schematic representation of experimental stimuli. A) Selected frames sequence depicting the biological motion pattern. B) Frames that, alternating, shape the rigid motion stimulus	79
Figure 4.4 A) Mean proportion of movements performed during stimulation and baseline trials (i.e. no stimulation) across fetuses. B) Looking times (ms) during stimulation and baseline trials across fetuses.	86
Figure 4.5 A) Mean proportion of fetal movements performed towards the stimuli and away from the stimuli, across fetuses. B) Looking times (ms) performed towards the stimuli and away from the stimuli, across fetuses.	86
Figure 4.6 A) Proportion of orienting movements across fetuses compared per condition. B) Looking times (ms) across fetuses compared per condition.	87

List of Tables

Table S1 Inclusion and exclusion criteria for eligibility.	36
Table S2 PubMed search strategy based on MeSH terms and keywords.....	37
Table S3 Studies assessing fetal movements based on outcome measure	45
Table 3.1 Precision, Recall, and F1 scores.....	61
Table 3.2 RMSE values.....	62
Table 3.3 F1 scores and RMSE values of the 250 frames subset.....	63
Figure 3.4 Example of trajectory and velocity profiles based on DeepLabCut predictions.....	25
Table 4.1 Experimental trial sequence	80

Chapter 1 - Introduction

1.1 The origin of fetal behavior

Prenatal behavior, defined as any observable movement, self-initiated by the fetus or externally triggered, has long held allure and mystery, to some extent due to its relative inaccessibility. Research on fetal behavioral development is still a relatively new enterprise. Systematic investigation of fetal behavior began in the late 1970s. As recently as the 1960s, many pediatricians and behavioral scientists still believed newborns to be a *tabula rasa* (James 1890), assuming that sensory and perceptual systems are inactive during the fetal period. The prenatal period was thus viewed as a “developmental limbo” (Hopkins & Johnson, 2005), probably due to the inability to directly observe and quantify fetal behavior during gestation. Research findings, mostly based on animal models (e.g., studies of embryonic and fetal behavior in chicks, rats, sheep, and primates), have progressively altered this conceptualization. It has been proposed that the fetus is vitally involved in its own development, adapting to the intrauterine environment by learning and preparing for the postnatal changes in its world (Lecanuet et al., 2013). The experimental study of developmental psychology began to address the potential for prenatal origins of developmental capacities. This was initially investigated by a set of studies providing indirect measures of fetal capacity (DeCasper & Spence, 1986; Schaal, 2000). The logic of these studies was to introduce an intervention (e.g., a sound) into the prenatal environment and then assess the behavior postnatally to determine if the change in the prenatal environment has had an effect (e.g., to test for preferential discrimination during the neonatal period). In the mid-to-late 1970s, the advent of ultrasound in obstetric practice revolutionized the field of obstetrics. Direct visualization allowed clinicians to see and describe the fetus and its behavior in real-time. Two-dimensional ultrasound (2D US) became the method of choice for characterization of behavior and variables used to assess that (e.g., eye and body movements). Fetal behavioral variables such as fetal body,

eye, breathing, and mouth movements and fetal heart rate gained increasing interest from clinicians who began to study them both isolated and in patterns. The first observations have identified the onset and characterized the maturation of spontaneous behaviors including heart rate, its variability and patterns, as well as body, limb, eye, breathing, and mouth movements. Over gestation, it has been observed increasing coordination among these behaviors resulting in organized behavioral states. Chapters of this dissertation are dedicated to the analysis of fetal body movements (Chapters 2 and 3) and eye movements (Chapter 4), with the ambition to develop new methodologies to better quantify such isolated variables and turn them into reliable data. Robust characterization of single variables of fetal behavior is the first essential step without which no reliable measurements would be possible. Larger studies will therefore be needed to fully understand how the application of new techniques can explain fetal behavior and neurodevelopment, both in normal and abnormal pregnancies.

1.2 What we currently know about fetal behavior (and brief remarks on the prenatal neuromotor development)

Fetal neuromotor development has been addressed from various perspectives and theories. An examination of all these theories as well as the emerging debate around cortical involvement in generating movements is beyond the scope of the present paragraph. Conversely, the purpose is to provide a brief overview of the emergence of fetal movements, supported by short notes on their neuroanatomical basis. The repertoire of fetal activities and functions during pregnancy increases as the central nervous system (CNS) develops. Understanding fetal behavior and neuromotor development could thus improve the assessment of the fetal neural system integrity. Systematic 2D US observations allowed the study of the motor response of healthy fetuses in their natural environment (de Vries & Fong, 2006), providing ample knowledge of the ontogeny of spontaneous behavior in the human embryo and fetus (de Vries & Fong, 2006; Visser et al., 1992). The initial studies using this approach were primarily observational.

In a nutshell, the earliest synapses in the spinal cord are detected between the 6th and 7th weeks of gestation (Okado, 1981). Simultaneously, embryonic movements (i.e., sideways bending) can be first detected at 7 weeks postmenstrual age. These first movements result from the neural activity of spinal motoneurons (Okado & Kojima, 1984). Simultaneously with the onset of spontaneous movements, the earliest motor reflex activity appears, indicating the existence of the first afferent-efferent circuits in the spinal cord (Okado, 1981). Then a set of specific movement patterns emerge, such that the fetus exhibits an impressive repertoire by 15 weeks of pregnancy, including general body movements (GMs), breathing, and head, limb, mouth, and eye movements. At 8-9 weeks, GMs are the earliest complex and well-organized movement patterns including the head, trunk, and limbs (de Vries et al., 1982). After the 10th week, the number, frequency, and diversity of fetal movements increase. GMs, which used to be slow and of limited amplitude, become more pronounced (Hadders-Algra, 2018). There is evidence from post-mortem studies demonstrating that by 12 weeks motor peripheral nerves in the upper limb have reached and innervated their muscle targets in human fetuses (Unver Dogan et al., 2009). From 12 weeks onwards, arm extension is frequently accompanied by finger extension. At the same age, the hand regularly touches the head, face, and sometimes the mouth. The neural infrastructure for sensory and proprioceptive feedback appears relatively early in gestation. Classical experiments in postmortem fetal tissue demonstrated reflexive responses to tactile cutaneous stimulation very early in development (Hakamada et al., 1988; Hooker, 1960), while muscle spindles mature between 20-30 weeks (Sarnat, 2003).

In the second trimester, fetal motor activity and behavior differ as a consequence of the development of neural connections, axons, synaptogenesis, and dendrite proliferation. Structures of the brainstem continue to mature, causing an increase in the complexity of fetal behavioral patterns and activities. Also, the second half of the pregnancy is characterized by a gradual organization of fetal movement patterns. The periods of fetal quiescence increase and the rest-activity cycles become recognizable. The brainstem, on the other side, gradually matures and begins to take over control of fetal movements and behavioral patterns (Joseph, 2000). A wide repertoire of fetal activity has been observed during the second trimester, including GMs, isolated movements of limbs, retroflexion and rotation of the

head, and facial movements like yawning, hiccuping, thumb sucking, swallowing, and mouthing (de Vries et al., 1982). For each pattern, the developmental trajectory has been studied in terms of age at emergence (i.e., the onset), range of occurrence, and change over time and in relation to other movement types. The first appearance of specific fetal movement patterns reported by (de Vries et al., 1982) is displayed in Figure 1.1. However, not all authors agree on the precise timing of these changes in occurrences. As a result of the reduction in intrauterine space, movements of the limbs tend to decrease during the second semester, ending up not far from the face.

During the third trimester of pregnancy, the CNS continues its development. The corticospinal tract (CST), originating principally in the primary motor cortex, contains axons that modulate the activation of alpha motor neurons being vital to the maturation of motor behavior (Sarnat, 2003). Unmyelinated CST axons reach lumbar regions of the spinal cord by 24 weeks, though maturational processes in humans are not complete until 2 years of age. Evidence for early CST functionality has been found using transcranial magnetic stimulation (TMS) of the motor cortex in ex utero preterm infants, where motor responses in the upper limbs are evoked from 26 weeks-equivalent age (Eyre et al., 2001). Regarding motor development, there is a decrease in the number of the GMs, and the head and the arm movements. On the other hand, these movements become more and more complex. In the third trimester, fetal eye movements are particularly interesting. Oculomotor control is likely dependent on the extraocular muscles, cranial nerves, and brainstem nuclei which develop at 7–9 weeks of gestation (Joseph, 2000). The first eye movements can be seen on 2D ultrasound at 14 weeks of gestation. At 26 weeks (post-conception), the fetus has eye lids partially open, by 28 weeks the eyes are wide open and at 28 to 31 weeks the pupils can constrict, dilate, and detect light (Donovan et al., 2020; Eswaran et al., 2004). Most studies on fetal eye movements are over 20 years old and were conducted once the efficacy of 2D US in the imaging of the fetal orbit in utero had been demonstrated. The aim of these studies was generally to identify (Okawa et al., 2017) and qualitatively classify fetal eye movements (Birnholtz, 1981) and their relevance to neurological abnormalities. There are very few studies looking at the fetal response to light stimulation such as flashlights applied directly to the maternal abdomen (Kiuchi et al., 2000), and simple patterned light stimuli (Donovan et al., 2020). These

stimulations produced increased eye movements, body movements, and fetal heart rate, indicating that fetuses respond to light stimulation.

More than 30 years ago, the recording of dynamic movement in fetuses became available with the introduction of four-dimensional ultrasonography (4D US). Observation of the fetus by 4D US has provided ample availability of data about fetal behavior from the first trimester through the entire fetal phase. The aim of these studies (see Chapter 2 for a review) was first to replicate findings of 2D US by assessing the quality and frequency of movement patterns, then to provide quantitative measurements of movements. However, a closer look at the literature highlights unaddressed challenges and limitations, raising concerns about the utility and credibility of 4D US studies of fetal movement.

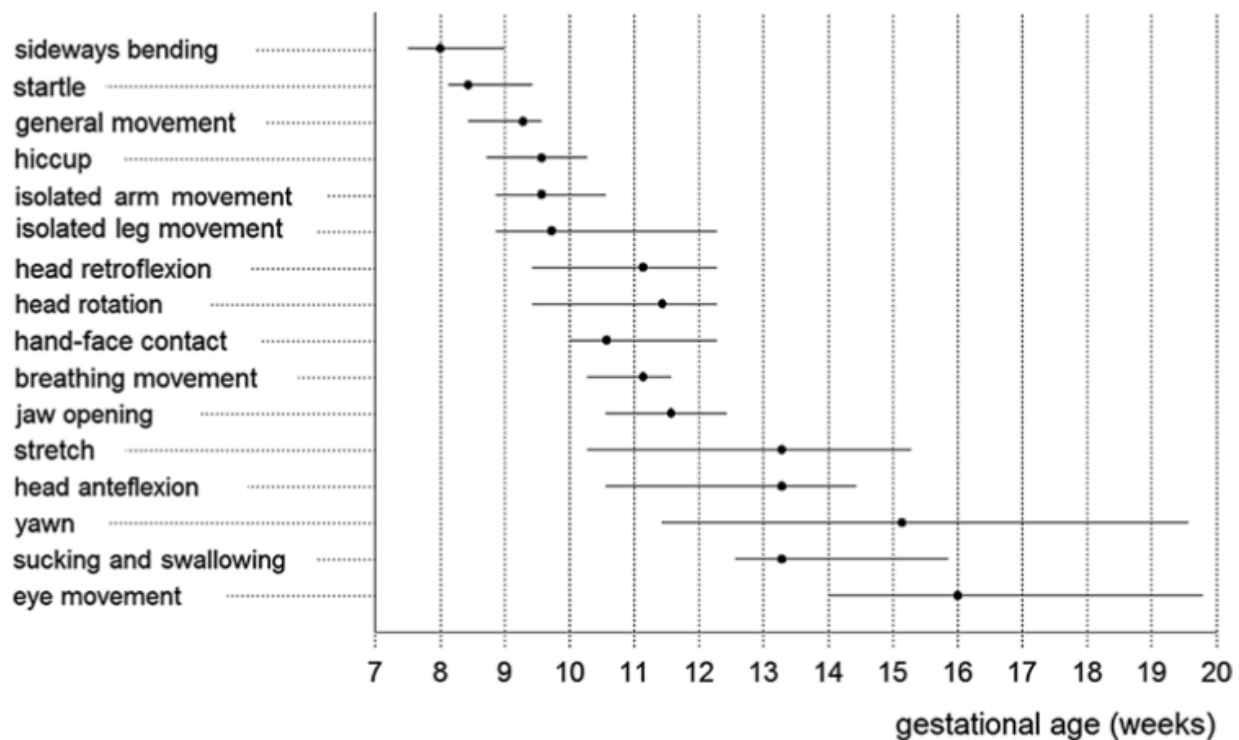


Figure 1.1 Emergence of specific fetal movement patterns with time presented as median and range of the first appearance during serial observations on individuals. Reproduced from (de Vries et al., 1982).

1.3 Why study fetal behavior?

Bringing together research on fetal brain and behavioral development is important for several reasons:

1. First, for the physiological contribution of early movements in promoting skeletal and muscular development. Motor activity is known to have beneficial effects on bone development in preclinical models and human preterm infants (Hagihara et al., 2009; Moyer-Mileur et al., 2008).

2. Fetal movements are spontaneous expressions of the central nervous system (CNS) and also contribute to shaping the CNS structures. Measures of the normal occurrence and development of fetal motor behavior could serve as early markers to monitor the integrity and development of the fetal CNS. This knowledge is essential to differentiate between the healthy fetus and the one that is (acutely) distressed, deteriorating, or has a chronic dysfunction. This leads to the potential ultimate goal of fetal behavior assessment which is to detect the effects of a particular antenatal disturbance and also to characterize these effects (Mulder & Visser, 2016). The complexity of fetal behavior assessment and the related challenges will be discussed in par. 1.4.

3. Fetal movements have the potential to play a key role in understanding fetal psychology. Early movements have been proposed to generate experiences for learning (Delafield-Butt & Gangopadhyay, 2013; Reid & Dunn, 2021; Reissland & Hopkins, 2010) and the fetus' ability to anticipate the sensory consequences of movements has been introduced (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007). Theoretical and methodological perspectives on the cognitive life of the human fetus are converging more and more towards the emerging field of fetal psychology. This will be discussed in par. 1.5.

4. Finally, fetal behavioral studies facilitate research linking disturbances in prenatal development to long-term outcomes, as part of the Fetal Programming hypothesis (Seckl & Holmes, 2007) and the Developmental Origins of Health and Disease (DOHaD) (Barker, 1990; Barker, 2004). The hypothesis states that, during critical or sensitive periods of development, a disturbance in environmental factors, such as

exposure to nutrient restriction, exposure to glucocorticoids or synthetic glucocorticoids, has an organizational effect on biological systems with a long development and/or intrinsic plasticity to react and adapt to environmental influence.

1.4 Reconceptualize fetal behavioral assessment

Impaired motor function is common to many pathological conditions in early life. The assessment of motor function has historically constituted an important component of the neurological examination in early infancy (Hielkema & Hadders-Algra, 2016) since abnormalities in posture, tone, reflexes, general movements, and gross and fine motor behavior has been reported to correlate with neuropathology (Einspieler et al., 1997; Prechtl & Einspieler, 1997). Infant behavior has been investigated for neurological evaluation through several assessment measures, each with its peculiarities. Some of these are the Apgar score (Watterberg et al., 2015), the Neonatal Behavioural Assessment Scale (NBAS) (Barlow et al., 2018), the NICU Network Neurobehavioral Scale (NNNS) (Salisbury et al., 2005) and the Newborn Behavioural Observations (NBO) (Barlow et al., 2018). The abundance of such measures and scales reflects the need to assess the neurobehavior early in infancy and, potentially, even prenatally.

Although there is no universally accepted procedure to test CNS functioning in the fetus to date, much has been achieved in fetal neurology over the last years (Mulder & Visser, 2016). The introduction of 2D US imaging made it possible to observe the fetus safely, directly, and repeatedly in its natural environment allowing the assessment of spontaneous fetal behavior. Following the assessment methods developed for preterm infants, full-term neonates and young infants, a few decades ago the urgent need for early information and feedback on CNS development in the prenatal stage became evident. Although an interactive fetal examination is not feasible, an assessment of spontaneous and reactive motor behavior may serve as a marker and predictor of neurological integrity (DiPietro, 2005).

The current focus of antenatal neurobehavioral assessment is to gather the information that reveals neural continuity from fetus to child. Such knowledge can provide basic scientific information regarding normal human ontogeny and identify antenatal factors that influence the trajectory of development. Thus, neurobehavioral assessment measures the performance along a continuum to distinguish among individuals within as well as beyond the normal range. The ultimate clinical application of fetal neurobehavioral assessment will be to identify functional characteristics of the fetus that predict a range of subsequent developmental dysfunction (DiPietro, 2005). Establishing this link will require demonstration of positive and negative predictability to outcomes significantly beyond the immediate perinatal period.

No unified neurobehavioral assessment method of the fetus currently exists to date. Instead, most investigators focus on one or more aspects of functioning within four general domains that include fetal heart rate, motor activity, behavioral state, and responsivity to stimulation. Fetal body movements give important information about the condition of the fetus, where, as a rule, the presence of normal movements is associated with a good fetal condition. Therefore, observations evaluate various aspects of movement, including tone, posture, reflexes, and activity level. Specifically, in evaluating fetal movements using 2D and 4D US, assessment tools have focused on amplitude, speed, complexity and fluidity (de Vries & Fong, 2006; de Vries et al., 1982; Salisbury et al., 2005; Sival, 1993), frequency of body movements occurrence within a given time window (Kurjak, Antsaklis, Stanojevic, & Porovic, 2017; Manning et al., 1990), and the application of concepts as the spatial use, timing, and force (Green & Sparling, 1993).

However, concerning the domain of fetal movement evaluation, most of these procedures show methodological limitations and bring out some challenges that should be addressed:

1. Most classifications of the motor activity do not describe unambiguously which body parts participated and whether a movement is performed in isolation or with ongoing activity elsewhere, leading to difficulties in comparing movement quantification across studies. This challenge applied to studies assessing movement both with 2D and 4D US. As for the latter, an extensive discussion is provided in Chapter 2, par. 2.1.

2. Traditionally, the coding of fetal movement has been done in real time by trained observers, or, more frequently, offline by recording the session for later review. In both cases, an accurate visualization requires a significant degree of expertise. Although detailed coding procedures are not always reported in fetal assessment descriptions, some minimum criteria should be met, such as the involvement of at least two independent coders and the quantification of inter-rater reliability. This would reduce the subjectivity introduced by observers when coding the movement under study.
3. Building a fetal assessment scale require rigorous psychometric development. Moreover, a sufficient degree of normative data available for each measure at different gestational ages drawn from representative populations is needed (DiPietro, 2005).
4. The use of the assessment tools is often limited to research purposes as they are too long for the medical assessments routine of fetal neurobehavioral development. As an example, the manual scoring of the 2D ultrasound scans within the Fetal Neurobehavior Coding System (FENS) introduced by Salisbury and colleagues (Salisbury et al., 2005) requires approximately 2-4 hours per tape.

With the availability of 4D US data at a high temporal resolution, a promising direction to improve the quantitative analysis of fetal movements is the application of deep learning and, in particular, deep neural networks. In a very near future, the use of automated tools will aid the assessment and measurement of fetal behavior by reducing the subjectivity introduced by observers in movement identification (thus overcoming the challenge n.2), but also eliminating the element of time required by manual coding (challenge n.4). Along with the adoption of (more) rigorous procedures for data collection and analysis, these technological advances will offer the opportunity to reconceptualize the domain of motor activity within the fetal neurobehavioral assessment methods. Such methods could benefit from a more objective and reliable characterization of fetal behavior. Chapter 3 of the thesis fits into this direction by developing a quantitative approach to make fine motion tracking of fetuses in the first and second trimesters.

Beyond spontaneous fetal movements, neurobehavioral assessment tools have been further investigated the fetal functioning focusing on the fetus' responsivity to stimulation. Documentation of fetal

response to stimulation can foster knowledge on the development of fetal sensory capacities and the rudiments of information processing. Traditionally, most interest has been dedicated to demonstrating fetal responsivity to vibroacoustic stimulations (Das et al., 2019; Tan et al., 2013), referring to stimuli activated on the maternal abdomen, usually above the fetal head. However, fetal responsiveness to light stimulation of the maternal abdomen has also been documented under various conditions (Eswaran et al., 2004; Kiuchi et al., 2000; Reid et al., 2017). Fetal responsivity to external light stimuli (including visual stimulations) is an important domain of fetal behavior having the potential to differentiate healthy and at-risk fetuses. Further investigations, by studying fetal responses in reaction to external stimuli, will need to demonstrate the added value these procedures can bring to the neurobehavior assessment. Chapter 4, by testing the feasibility of a novel paradigm to deliver light stimuli in utero, aims to be part of this research line.

Over the years, there has been growing interest in the assessment of the integrity and activity of the fetal CNS. During the postnatal life we can actually ‘test’ the CNS, rather than observe it, but prenatally that remain a challenge. Behavior is the most important output of neural activity and quantifying behavior is crucial for understanding how the brain works and develops. Studies examining prospective motor control in fetuses (Castiello et al., 2010; Myowa-Yamakoshi & Takeshita, 2006; Reissland et al., 2014; Zoia et al., 2007), as well as further replication of their results, could provide insight into neurodevelopmental disorders in which the development of anticipations skills appears delayed (Sinha et al., 2014; Trevarthen & Delafield-Butt, 2013). By providing an effective tool to extract movement parameters, Chapter 3 stands as an essential first step in performing a careful analysis of fetal movement kinematics. An automated and reliable assessment of fetal behavior has thus the potential to detect abnormal motor patterns of development in utero.

1.5 Towards a fetal psychology and beyond

Previous paragraphs pointed out that fetal movements serve important biological mechanisms and are crucial for the fetal physical or physiological well-being. Besides these established functions, fetal movements might play a key role in understanding fetal psychology (Delafield-Butt & Gangopadhyay, 2013; Reid & Dunn, 2021; Reissland & Hopkins, 2010). Examining human development starting from life in the womb brings new perspectives and challenges to the understanding of psychological development, some of which will be considered in this thesis.

As gestation proceeds, the fetus is exposed to an increasing number of local contacts and general displacements. These contacts can be with the uterine wall, as well as with the fetus' own body (e.g., hand-body, hand-face, hand-mouth contacts) and with the co-twin in twin pregnancy. Late in the first trimester of gestation, the fetus performs isolated movements of head and limbs, showing a differential control of elements of its skeletomusculature. With isolated movements, fetus seems to increasingly prefer those parts that are richly innervated. From 10-12 weeks, face contacts are highly frequent, which is interesting knowing that the trigeminal innervating the face is an important source of tactile and proprioceptive sensations (Fagard et al., 2018). These hand-to-face movements, touching sensitive innervated regions, create an autostimulatory feedback loop. The contact between the fingers and head give simultaneously a proprioceptive response in the resistance met at contact, the sensation of touch in the fingers, and the sensation of touch in the innervated region (Delafield-Butt & Gangopadhyay, 2013). As pregnancy progresses, arm movements aim increasingly toward the mouth (Piontelli, 2010). This might indicate that fetuses progressively select movements inducing interesting sensory feedback. Learning this particular temporal structure may serve as a basis for handling causality and prediction.

The scenario of a fetal awareness of the consequences of self-generated actions (Ciaunica et al., 2021) has opened the doors to a new line of studies investigating the prenatal onset of motor control. For instance, two studies have addressed the issue in second and third-trimester fetuses by examining the co-occurrence of hand-to-face movements and mouth movements. They found that fetuses anticipate their

movement toward the mouth by opening the mouth before the hand arrives (Myowa-Yamakoshi & Takeshita, 2006; Reissland et al., 2014). This anticipatory behavior seems to increase in frequency as gestation progresses (Reissland et al., 2014), as a result of brain maturation and increased level of motor control. Findings have been interpreted as evidence that predictive processes might already be operating in fetuses. Other studies introduced fetal kinematics as a tool to determine whether fetuses calibrate prospectively their reaching as a function of the target (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007). The reaching of singleton fetuses is found to be straighter and more directly aimed toward the target by 22 weeks. Moreover, at 22 weeks they found longer movement duration and deceleration time for movements towards the eyes rather than the mouth (Zoia et al., 2013; Zoia et al., 2007). According to the authors, acceleration and deceleration phases were planned depending on the size and/or delicacy of the target. A similar approach was extended to twin fetuses, where calibration effects were found even earlier, at 14 and 18 gestational weeks, as well as longer movement duration and deceleration time for movements towards the co-twin, compared with self-directed movements (Castiello et al., 2010). All of these studies focused on a key component of motor control, namely the early potential of the fetus to anticipate the sensory consequences of a movement and use them to plan an action according to the nature of the target. As in the postnatal life, differential movement patterns of fetuses observed for hand to eye, hand to mouth movement, and movement towards the co-twin might thus indicate that information about different sensations obtained by the target are used to adjust the approach of the hand (Hofsten, 2009). The huge ambition of these studies is a significant starting point to shed light on fundamental questions of prenatal life and human nervous system development. However, procedures used to compute these movement parameters might be not sufficient to evaluate the accuracy and precision of these measures. As an example, the slow frame rate of US imaging in the above studies limits the range of movement frequencies that can be measured. This technological challenge is addressed extensively in Chapter 2.

Considering the importance of accurate quantification of prenatal motor control, the need to replicate these findings by collecting US scans at higher temporal resolutions and standardized reporting

practices is even more pressing. Chapter 3 of the thesis fits into this research line and lays the groundwork for further kinematic studies, introducing new opportunities to conventional ways of quantifying fetal movements.

While most of the above-mentioned examinations focus on spontaneous fetal movements, a set of studies investigated whether and to what extent fetuses respond to maternal stimuli. Marx and Nagy (Marx & Nagy, 2015, 2017) reported that 20- to 35-week-old fetuses respond specifically to their mother's stimuli. Fetuses increased head, arm, and mouth movements in response to the touch of the maternal abdomen and decrease arm and head movements to maternal voice. This approach was then expanded to different types of abdominal touch, showing differential fetal responses to mother's touch, father's touch and stranger's touch (Marx & Nagy, 2017). Other studies investigated the fetal sensitivity and behavioral responses to maternal acoustic stimulation during the third trimester (DeCasper et al., 1994; Ferrari et al., 2016; Hepper et al., 1993; Kisilevsky et al., 2003). Ferrari and colleagues (Ferrari et al., 2016) using 2D US observed more mouth openings when the mother sang the syllable 'LA' in a nursery rhyme than in any other stimulation (chewing, yawning, etc). In a similar study, fetuses reacted with a specific mouth movement that mimics the sound heard, upon hearing the repeated sound 'MA' (Reissland et al., 2016). Overall, results suggest that maternal touch and voice might be powerful stimuli, producing a range of fetal behavioral responses. However, it is unclear whether the effect found in these studies is a response to emotional arousal or a reflex-like response (Ferrari et al., 2016). A possible way to test this might be to analyze fetal heart rate variations in response to the mother or other individuals speaking with or touching the fetus. A differential heart rate response to the mother versus a stranger's voice/touch could provide evidence of the potential coupling between behavioral outcomes and emotional arousal involving the autonomic system.

Two studies (Reid et al., 2017; Reissland et al., 2020) assessed how the frequency of fetal head movements in third-trimester fetuses changes in response to visual stimulations and reported more fetal head turns in the direction of face-like stimuli compared to non-face-like control light stimuli. Research with neonates has consistently demonstrated that newborns show a preference for highly schematized face-like configurations, suggesting an innate preference. These two studies aimed to investigate whether the

acquisition of social preferences and expertise can be detected as early as in fetuses. In line with this approach, Chapter 4 develop a visual stimulation paradigm with the ultimate goal to study biological motion processing in fetuses. A spontaneous preference for looking at biological motion is observed from the very first days of life but no evidence has been reported before the birth. Moreover, evidence assigns visual preference for biological motion a critical role in many aspects of social and communicative development (Pavlova, 2011). Therefore, detection of the fetal ability to process biological motion could add important information about neural continuity from fetus to the child, providing useful insight into the characterization of normal human ontogeny.

The emerging field of fetal psychology is gaining increasing appeal, although more empirical work is needed before we can understand the complex interactions between the fetus and the external environment in which the fetus develops and grows. One future direction is related to the emergence of intentionality (Delafield-Butt & Gangopadhyay, 2013). Fetal movements have often been interpreted as a way for the fetus to practice and exercise its emerging motor system, participating in its maturation. General movements and isolated movements allow the fetus to explore its body as well as its environment. Motor signals and self-touch behavior, resulting in touch exploration of the fetus' own body, have been proposed to give rise to a basic form of body schema (Haggard & Wolpert, 2005; Kuniyoshi, 2019), defined as a sensorimotor representation of self. Instead, within this perspective, non-self objects (i.e. the uterine wall and umbilical cord) are assigned separate sensorimotor representations that might become the basis for object representations. Self- and environment exploratory movements can be viewed as a way to perceive and explore themselves, the world, and their relation to the world via repeated trials (Ciaunica et al., 2021). These embodied experiences, despite the small action-space within the uterus, have been proposed to constitute the roots of content-directedness (Ciaunica et al., 2021; Delafield-Butt & Gangopadhyay, 2013).

Indirect support is provided by kinematic studies (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007) suggesting a sensory knowledge of the fetus of its environment. This was supported by reporting the fetus' early ability in planning and calibrating movements depending on the target. Challenges and limitations when conducting fetal kinematic studies will be discussed in Chapter 2.

Nevertheless, Chapter 3 aims to overcome some of these limitations by providing reliable movement parameters of fetuses recorded with 4D US. Chapter 3 aspires to stimulate future research with reproducible methods that may add new findings to the debate on prospective fetus' awareness of the sensory consequences of its movements and fetal consciousness. The recognition of the philosophical implications of fetal movements on the nature of consciousness further increases interest in this topic.

1.6 Difficulty of measuring fetal behavior

The number of challenges involved in US assessment of fetal behavior, as well as over-simplification and over-interpretation of the data, might have led to a certain reticence to embrace fetal behavior research (Hepper, 1996). The methodology used to monitor fetal behavior (both 2D and 4D US) may be a cause of concern, for its perceived subjectivity and time-consuming nature. Different observers may interpret differently and introduce subjectively into their observations, thereby reducing the usefulness of fetal behavior studies. Indeed, most studies define the behaviors under study, however, definitions are rarely turned into operational measures that can be unambiguously applied to specific movement events. Moreover, the documentation of behavior is very time-consuming. Studying fetal behavior requires a lot of experience, prolonged observations, big numbers, off-line analyses performed by more than one operator, and extensive data elaboration (Piontelli, 2015).

Measuring fetal behavior through 4D ultrasound technology raises further and novel problems compared to the 2D US. The resolution of the 4D scans, for example, is relatively poor. The surface rendering quality of 4D ultrasound is dependent on temporal resolution: the higher the sampling rate, the lower the quality of the rendered image. Also, the limited field of view affects the capability of visualizing the whole fetus beyond 20 weeks of gestation.

Beyond these common issues when conducting fetal behavioral research, a significant number of difficulties arise when exploring visual processing in fetuses (Reid & Dunn, 2021). For example, the attrition rate of fetuses recruited in most studies is very high. Typically, only approximately 50% of the

sample provides data that can be used in the analysis (Dunn et al., 2015). In studies performed in our units, this rate was even lower. Only about 20% of pregnant women giving the consensus to participate were then actually involved in our visual stimulation experiment. This was mainly a consequence of the difficulty of collecting data without being able to modify the fetus' initial position. When studying fetal response to light stimulation, the fetus may be positioned in such a way that no stimulus can be presented. This happens, for example, when the placenta or hand/s occlude the fetal face, when the face of the fetus is directed towards the mother's back and when the fetus is judged to be asleep. To overcome the intrinsic barriers to the delivery of visual stimuli in the uterus that have hindered progress to date, a great effort in the optimization of paradigms and stimuli is required.

1.7 Thesis overview

In this chapter, I have outlined how the ontogeny of human behavior and movement have been studied through gestation, by providing brief remarks on prenatal neuromotor development. After the advent of ultrasound, fetal behavior was directly assessed through 2D real-time ultrasound. Later, a continuous visualization of fetal movements via 4D ultrasound became available. However, the impact of 4D US on obstetric medicine and research has been limited. Despite multiple barriers, the importance of studying fetal behavior is growing. I have pointed out the interest in fetal movement and behavior from different domains: clinical, psychological, and philosophical. Then, I have highlighted some of the common challenges encountered when conducting fetal research with ultrasound, followed by several practical difficulties when exploring visual processing in fetuses. The primary aim of this thesis is to develop a new quantitative framework to investigate fetal behavior and movement. This was done by striving to overcome some of the challenges and limitations in quantifying and interpreting US fetal movement data one is confronted with when reviewing the literature. Here, the fetal behavioral development is investigated by i) tracking fetal movement assessed with 4D US, and ii) evaluating fetal responsivity to visual stimulation assessed with

2D US. To do so, I employed different techniques, including ultrasound, deep neural network, motion capture system, and attentional measures. Below I highlight the specific questions addressed in each chapter:

Chapter 2. Does 4D ultrasound significantly improve fetal movement research?

Chapter 2 reviews what is currently known on fetal movements assessed using 4D US across gestation. The chapter first provides a brief overview of the challenges encountered in tracking fetal movements using 4D US. Then, based on the results of a systematic review, the chapter examines how the various authors report and handle the issues that arise from such challenges. This literature review concludes that methodological and technological advances are needed for 4D US characterization of fetal behavior to qualify for scientific and clinical use and point to promising future directions, including the application of deep learning.

Chapter 3. Can we apply deep neural networks to track fetal movements?

Chapter 3 reports on a study aimed at generating the first objectively annotated dataset of fetal movements. To our best knowledge, there is currently no annotated dataset that can be used to study fetal movements. To develop such a dataset, we applied a deep convolutional network (DeepLabCut) to 4D US fetal scans. 4D US scans were first manually annotated by three independent labelers. Annotations were then combined to create a consensus annotation and to evaluate the performance of the algorithm against humans in tracking the position of user-defined body parts. DeepLabCut performs the task at a human level, suggesting that the algorithm succeeds in tracking fetal movement from high-resolution 4D US images. The trained network will be released to the scientific community, this will allow the user to analyze new videos and extract reliable movement parameters bypassing the most time-consuming phase of manual labeling.

Chapter 4. Do fetuses show a preference for biological movement?

Chapter 4 reports on a study designed to develop a new methodology to investigate fetal eye movements in third-trimester fetuses through the detection of the eye lens in 2D US scans. The ultimate ambition is to use this paradigm to investigate whether a preference for biological movement already exists before birth. In this pilot study, complex visual stimuli were projected to the maternal abdomen using a LEDs stimulator.

Stimuli configurations displayed a biological motion pattern and a rigid motion pattern, alternately. Fetal reactions (i.e. head and eye movements) to stimuli were observed in fetuses by simultaneously recording movements using 2D US. The chapter reports on the results collected on 7 fetuses, suggesting that our stimulation protocol is effective in eliciting a response from fetuses and can be used to test specific predictions deriving from the hypothesis of a fetal preference for biological movement.

Finally, Chapter 5 summarizes the findings of my thesis and discusses open questions and future directions.

Chapter 2 - Why 4D ultrasound has not (yet) revolutionized fetal movement research

The results of Chapter 2 were published in *Ultrasound in Obstetrics and Gynecology*:

Pretti N., Paladini D., Panzeri S., Becchio C. (2022). Why 4D ultrasound has not (yet) revolutionized fetal movement research, *Ultrasound in Obstetrics and Gynecology* 59 (5). DOI: 10.1002/uog.24757

Behavior is the most important output of brain activity and hence, by studying behavior, it is possible to make inferences about the functioning of the brain. The introduction of four-dimensional ultrasonography (4D US) into clinical practice was expected to provide unprecedented opportunities to study fetal behavior (Piontelli, 2015; Reissland & Kisilevsky, 2016), assess fetal well-being, and determine fetal brain function. Despite these expectations and more than 30 years of research, the impact of 4D US on obstetric medicine and research has been limited, raising concerns about the utility and credibility of 4D US studies of fetal movement. This opinion considers possible reasons for this. We first provide a brief overview of the challenges encountered in tracking fetal movements using 4D US. Then, based on the results of a systematic review (available as Supporting Information), we examine how the various authors report and handle the issues that arise from such challenges. We conclude that methodological and technological advances are needed for 4D US characterization of fetal behavior to qualify for scientific and clinical use and point to promising future directions.

2.1 Challenges involved in 4D US assessment of fetal behavior

4D US offers several advantages over two-dimensional ultrasound (2D US) (Merz, 2005; Pretorius & Nelson, 1995). Most importantly for the study of fetal behavior, 4D US provides continuous visualization of fetal movements (de Vries & Fong, 2006). With 2D US, only a single cross-sectional plane is imaged at a time and the visualization of movement, such as upper-limb movement toward the fetal face, can be achieved only by moving the transducer and changing the angle of insonation by hand, to ‘chase’ the moving limb (making impossible the assessment of movement velocity). 4DUS allows movement to be visualized continuously within a three-dimensional (3D) region of interest (ROI). However, the assessment of fetal movements remains difficult. Herein, we review briefly the unique challenges involved in identifying, classifying and measuring fetal movements with 4D US.

Identifying fetal movements. The first challenge is to segment movements, that is, to identify segments that contain movements and differentiate them from non-movement segments. Deciding when a movement begins and ends might sound straightforward but is not. In postnatal studies, movement segmentation can benefit from objective measures of position or position derivatives, mainly velocity, or higher derivatives, such as acceleration or jerk (Georgopoulos et al., 1982; Moran & Schwartz, 1999; Roby-Brami et al., 2000; Scott et al., 2001). In prenatal studies, none of these measures is available. Segmentation is performed manually by evaluating the change in position from one frame to the next of a user-defined key point (e.g. wrist). This procedure is methodologically weak as it is highly subjective and does not guarantee consistent segmentation across users. Moreover, it is difficult to determine the time of movement onset and offset when the body key point is not visible, for example, due to shadowing by other body parts or restriction in the field of view (Figure 2.1). Finally, with this procedure, the sampling of behavior is limited intrinsically by the frame rate of 4D US. With a frame rate of 4 Hz, for example, the finest achievable temporal resolution, defined as the interval between two separable events, is 250 ms (Figure 2.2). Whilst signal processing techniques (such as linear interpolation) may be used for resolution enhancement, to our knowledge, no such techniques have been used or validated for 4D US. Since movement onset and offset

are also used to assess other events, such as time to peak acceleration, time to peak velocity and deceleration time, these limitations also apply to measurement of fetal movements (see below).

Classifying fetal movements. Once movement segments have been identified, a second challenge is to classify them, that is, to assign movement segments to different movement categories. Most studies do so by defining the behaviors under study. However, definitions rarely reflect operational measures that can be applied unambiguously to specific movement events. For example, general fetal movements (de Vries et al., 1982) have been defined as a ‘variable sequence of arm, leg, neck, and trunk movements’ that ‘wax and wane in intensity, force, and speed, and they have a gradual beginning and end” (Prechtl & Einspieler, 1997). How should changes in intensity, force, and speed be determined? How should the gradual beginning and end of a burst of general movement be determined (Ten Hof et al., 1999)? In the absence of adequate and unambiguous operational definitions of the variables to be measured, different human coders might interpret the same observation differently, leading to inconsistency in classification of movements.

Measuring fetal movements. A third challenge relates to measuring fetal movements. In some motor-control studies, humans (and animals employed in similar studies) are fitted with reflective markers to assist with computer-based motion tracking (Ansuini et al., 2015). Alternatively, markerless systems combine multiple camera views with sophisticated feature-extraction algorithms to compute the motion of individual body segments without using markers (Mündermann et al., 2006). Neither of these approaches is applicable to fetal movements. Studies reporting on the kinematics of fetal movements extract motion parameters from manual digitization of 3D surface-rendered images. With this method, out-of-plane motion of body segments can corrupt in-plane displacement measurements and derivatives, making the estimation and comparison of movement profiles difficult across movements with varying out-of-plane components. Moreover, the restricted field-of-view and the relatively slow frame rate of 4D US limit the range of displacements that can be tracked and the range of movement frequencies that can be measured. Based on the Nyquist-Shannon theorem, which establishes a fundamental limit on the resolution afforded by a given frame rate, a sampling rate of 4 Hz, as in the above example, is sufficient for capturing movements with

frequency components of 2 Hz or less (Nyquist, 1928; Shannon, 1949), but not movements with higher frequency components (such as rapid eye movements, startles or fetal seizures).

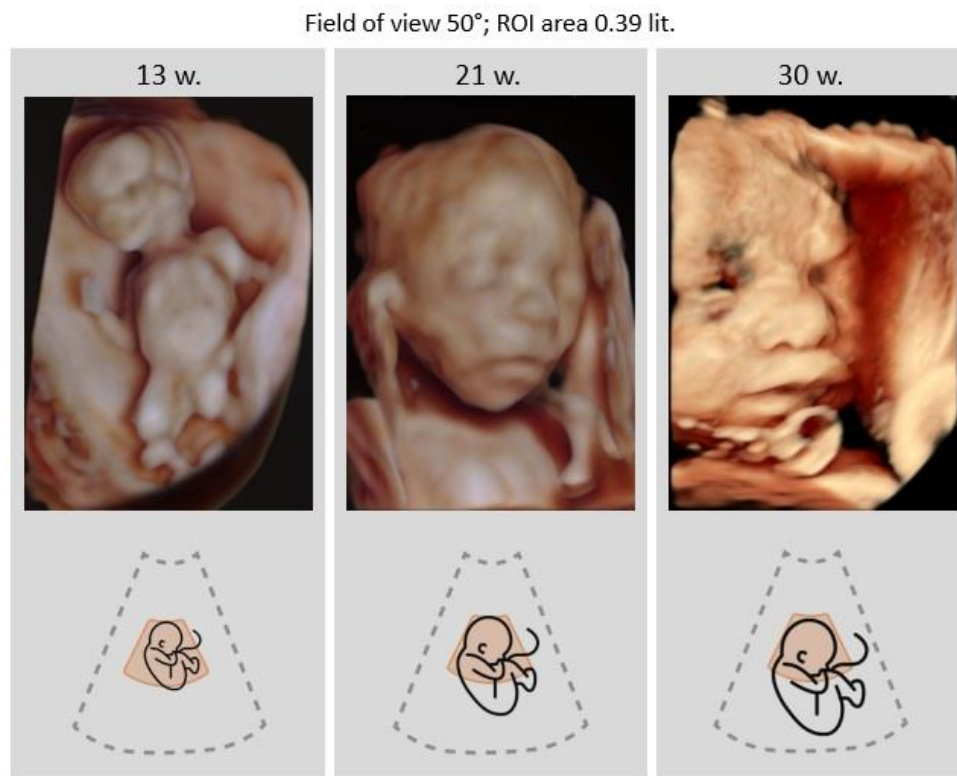


Figure 2.1 Impact of fetal growth on visualizing fetal movements at a fixed ROI volume. Three-dimensional scans of fetuses at 13, 21, and 30 weeks of gestation (upper row) were obtained using a Voluson E10 (GE Healthcare, Zipf, Austria) ultrasound system equipped with a RM7C volumetric transducer. Images were acquired keeping the volume angle of the field of view fixed at 50°, then adapting the ROI area to obtain a maximum volume of 0.39 lit. Schematic diagrams (lower row) illustrate how the angle of the field-of-view impacts the capability of visualizing fetal movements at different gestational ages. Due to the limited field-of-view (shaded area), only a partial view of the fetus is possible beyond 20 weeks of gestation.

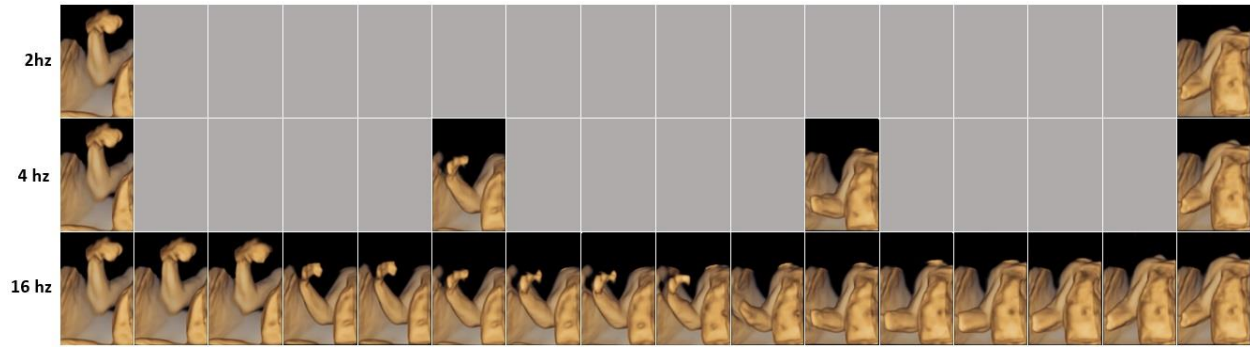


Figure 2.2 Impact of frame rate on visualization of fetal movements. Isolated arm movement (extension) acquired at 20 weeks of gestation using a Voluson E10 (GE Healthcare, Zipf, Austria) ultrasound system equipped with an eM6CG3 volumetric transducer and with frame rate set at 16 Hz. The bottom row displays the movement sequence at 16 frames/s. In the top and middle rows, the same sequence is depicted artificially down-sampled to 2 frames/s and 4 frames/s, to illustrate the impact of lowering the frame rate to 2Hz and 4 Hz, respectively.

2.2 Systematic review of 4D US fetal movement studies

These challenges raise critical methodological and interpretative concerns. To assess the impact of such concerns on fetal-movement research, we conducted a systematic review of 4D-US studies reporting on fetal head, limb and general movements in singletons and twins. The detailed results of our literature search and screening process are provided as Supporting Information (Table S3), along with a narrative synthesis of the 56 studies reviewed. Herein, we discuss how the reviewed studies reported and handled methodological and interpretative concerns in tracking fetal movements with 4DUS. We address these concerns in three clusters of studies, aggregated on the basis of outcome measures: (i) studies assessing the frequency of fetal movements ($n = 30$); (ii) studies assessing the quality of fetal movements ($n = 22$); (iii) studies assessing the kinematics of fetal movements ($n = 4$).

2.2.1 Assessing the frequency of fetal movements

We reviewed 30 studies assessing the frequency of fetal-movement patterns across trimesters in singleton and twin fetuses (Table S3). The scanning period in these studies varied greatly, from 2 to 60 minutes, with most examinations conducted over a period of 15 - 30 minutes (AboEllail et al., 2018; Andonotopo & Kurjak, 2006; Andonotopo et al., 2005; Degani et al., 2009; Hata et al., 2011; Hata et al., 2012; Kim et al., 2010; Kurjak et al., 2006; Kurjak et al., 2003; Kurjak et al., 2004; Kurjak et al., 2005; Kurjak, Talic, Stanojevic, et al., 2013; Lebit & Vladareanu, 2011; Marx & Nagy, 2015, 2017; Mori et al., 2018; Myowa-Yamakoshi & Takeshita, 2006; Prats et al., 2012; Reissland et al., 2014; Reissland, Francis, et al., 2015; Sajapala et al., 2017; Sasaki et al., 2010; Yigiter & Kavak, 2006). Of the 30 studies, only 15 (50%) reported frame rate. Seven of these studies (Hata et al., 2010; Hata et al., 2011; Hata et al., 2012; Marx & Nagy, 2015; Myowa-Yamakoshi & Takeshita, 2006; Reissland, Francis, et al., 2015; Sasaki et al., 2010) recorded fetal movements with a frame rate of 2 - 4 Hz. In the remaining eight studies (Andonotopo & Kurjak, 2006; Andonotopo et al., 2005; Kuno et al., 2001; Kurjak et al., 2006; Kurjak et al., 2003; Kurjak et al., 2004; Kurjak et al., 2005; Yigiter & Kavak, 2006), the volume was scanned automatically every 2 s, which corresponds to a frame rate as low as 0.5 Hz. Only one study (Reissland, Francis, et al., 2015) reported field angle. As frame rate and field angle have a decisive impact on movement segmentation, this severely limits the possibility to compare frequencies across studies and gestational ages. For example, a higher frequency of general movements in the first trimester could simply reflect the difficulty of visualizing whole-body movements by 4D US after 20 weeks (Figure 1). Also, the frequencies used in some studies may not have been sufficient to satisfy the Nyquist-Shannon sampling rate condition for capturing faster motions (Nyquist, 1928; Shannon, 1949).

Furthermore, coding procedures were also heterogeneous across studies. Eleven (Andonotopo & Kurjak, 2006; Kim et al., 2010; Kuno et al., 2001; Kurjak et al., 2006; Kurjak et al., 2003; Kurjak et al., 2004; Kurjak et al., 2005; Kurjak et al., 2002; Lebit & Vladareanu, 2011; Ogo et al., 2019; Yigiter & Kavak, 2006) of the 30 studies did not specify the number of coders involved, seven studies (AboEllail et al., 2018;

Degani et al., 2009; Hata et al., 2010; Hata et al., 2011; Hata et al., 2012; Mori et al., 2018; Sasaki et al., 2010) reported coding by a single rater and only 12 studies (Andonotopo et al., 2005; Kurjak, Talic, Stanojevic, et al., 2013; Marx & Nagy, 2015, 2017; Prats et al., 2012; Reid et al., 2017; Reissland, Aydin, et al., 2015; Reissland et al., 2014; Reissland et al., 2020; Sajapala et al., 2017) used coding by two raters. In these 12 studies, the inter-rater reliability was assessed using Cohen's kappa coefficient (Marx & Nagy, 2015, 2017; Myowa-Yamakoshi & Takeshita, 2006; Reid et al., 2017; Reissland, Aydin, et al., 2015; Reissland et al., 2014; Reissland, Francis, et al., 2015; Reissland et al., 2020), or other correlation coefficients (Sajapala et al., 2017). Two studies (AboEllail et al., 2018; Mori et al., 2018) reported only 'good interclass correlation coefficients and interobserver agreements' without specifying exact values. Others evaluated interrater reliability using a coefficient of variation (Kurjak, Talic, Stanojevic, et al., 2013; Prats et al., 2012) or a consensus reading between investigators that was not specified further (Andonotopo et al., 2005). Of the 12 studies, only four (Marx & Nagy, 2015; Myowa-Yamakoshi & Takeshita, 2006; Reid et al., 2017; Reissland, Aydin, et al., 2015) reported at least one coder to be blind to the experimental condition/hypothesis.

2.2.2 Assessing the quality of fetal movements

Twenty-two papers assessed the quality of fetal movements (Table S3), all using the Kurjak Antenatal Neurodevelopmental Test (KANET) (Kurjak et al., 2010). This scoring system comprises assessment of the following parameters: isolated head anteflexion, cranial sutures and head circumference, isolated eye blinking, facial alterations, mouth opening, isolated leg movement, isolated hand movement, hand-to-face movements, finger movements and thumb position, and gestalt perception of general movements. Each parameter is assigned a score of 0 (absent/abnormal) to 2 (normal) (Kurjak et al., 2010), or 1 to 3 (Kurjak et al., 2008), based on visual inspection. Fetal motor behavior is classified as abnormal if total score is ≤ 5 , mildly or moderately abnormal if it is between 6 and 13, and normal if it is ≥ 14 .

All 22 studies were based on 4D US examinations carried out over a period of 15 - 30 min. Despite KANET recommends a frame rate of at least 24 volumes/s (i.e., a frame rate of 24Hz) (Kurjak, Antsaklis, Stanojevic, Vladareanu, et al., 2017), only one (Hata, Kanenishi, Mori, AboEllail, Hanaoka, Koyano, Kato, Kusaka, et al., 2019) of the 22 studies reported the acquisition frame rate, and this was, regrettably, critically low (1-3 Hz). This raises serious concerns regarding the reliability of the reported results. Also, operational definitions and coding procedures were not specified except in two studies (Antsaklis et al., 2017; Athanasiadis et al., 2013), which reported that US scans were performed by two examiners but did not specify whether (and, if so, how) the two examiners scored the videos. This also gives cause for concern, because, even if they are given the same general descriptors of a behavior to be observed, different observers may interpret it differently and introduce unnecessary variability into the observations. None of the studies reported inter-rater reliability.

2.2.3 Assessing the kinematics of fetal movements

Four papers (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007) assessed the kinematics of fetal arm movements to determine whether early in the second trimester of pregnancy (14-22 weeks) fetuses calibrate prospectively their reaching as a function of the target (mouth region, eye region, co-twin, no target) (Becchio et al., 2018; Rosenbaum et al., 2012). In these studies, examinations were conducted over a period of 20 min, with the ultrasound probe set to 4 Hz, capturing the torso, arm and head of the fetus. Extracted frames were manually digitized and wrist displacement was computed with reference to the fetus body axes, using the intra-ocular distance as measurement unit. One study (Zoia et al., 2013) specified that a target was considered to have been reached each time the hand stopped on the target for at least two consecutive frames.

In all four studies, movements were classified by at least two experts, with high inter-rater reliability reported by two studies (Castiello et al., 2010; Zoia et al., 2013). One study (Castiello et al., 2010) reported both experts being blind to the experimental conditions. The wrist marker was used to compute arm velocity

(displacement derivative) data and determine movement time (in ms), computed as the time at which the wrist first began to move until speed was close to zero. Additionally, studies reported time to peak velocity (as %), defined as the time from movement onset to peak velocity in relation to the total movement time (Parma et al., 2017; Zoia et al., 2007) and deceleration time (as %), defined as the time spent from the peak velocity to the end of movement in relation to the total movement time (Castiello et al., 2010; Zoia et al., 2007). However, the procedures used to compute these kinematic parameters were not reported. For example, in the replication dataset included in one study (Parma et al., 2017), movement time ranged, across trials and fetuses, between 1054 and 3321 ms. The procedure used to determine movement time with a precision of milliseconds from a recording with a frame rate of 4 Hz was not specified. Also, potential confounders introduced by varying out-of-plane components were not quantified, contributing to the difficulty of evaluating the accuracy and precision of these measures.

2.3 Future is now

Major scientific breakthroughs are often achieved when old assumptions are challenged. Our systematic review highlights largely unaddressed methodological and interpretative concerns regarding the use of 4D US to characterize fetal behavior that should and can be addressed. One prominent concern arises from the low temporal resolution of 4D US. Better fetal-movement research requires higher temporal resolution. This can now be achieved (to some extent) thanks to the improved frame rate of the new matrix transducers. For example, in studies currently performed in our units, frame rates of 13-16 Hz are achieved reproducibly over predefined ROIs. With the availability of 4D US data with higher frame rates, a promising direction to improve the quantitative analysis of fetal movements will be the application of deep learning and, in particular, deep neural networks. In recent years, powerful deep neural network tools designed specifically to aid the measurement of behavior have come to fruition (Mathis et al., 2018). These tools enable the relatively straight-forward creation of deep networks highly specialized for a specific task. Current applications range from tracking mice in open field mazes to whole-body movements in flies and pose

estimation in human babies. In the future, we envisage networks trained on high-frame-rate 4D US data to track fetal movement and behavior. The introduction of automated tools will eliminate not only the element of time required by manual coding, but also that of subjectivity. The majority of measurements used in fetal movement studies appear to be highly subjective, as observers code the movement under study (Hepper, 1996). Automated pose estimation of fetal movement would remove the problem of subjectivity. These advances, along with the adoption of (more) rigorous procedures for data collection and analysis, will offer researchers in the field the opportunity to begin new, more objective and reliable, fetal movement research. By providing direct access to the output of brain activity - behavior - this research has a potentially important role in assessing fetal neurodevelopment and wellbeing.

2.4 Supporting Information

Systematic Review of the literature on fetal movements

The following sections report the Methods and the Results of a systematic review of 4D US studies of fetal head, limb and general movements.

2.4.1 Methods

2.4.1.1 Protocol and registration

This systematic review was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Liberati et al., 2009). Methods were registered on the Prospero International Prospective Register of Systematic Reviews (N. CRD42020154341).

2.4.1.2 Eligibility criteria

We selected papers reporting on fetal movements of the head, limbs and general movements recorded in singletons and twins using 4D US. We considered both prospective and retrospective cohort studies. References of selected papers and reviews related to the topic were examined for relevant citations. Table S1 presents the eligibility criteria for inclusion in this systematic review.

Table S1. Inclusion and exclusion criteria for eligibility

Inclusion criteria:

1. Studies using 4D US
2. Studies focusing on head, limbs and general movements
3. Studies involving fetuses in the first, second and third trimester of pregnancy

Exclusion criteria:

1. Review articles, single case reports and book chapters
 2. Unpublished data and grey literature
 3. Not peer-reviewed articles
 4. Articles not written in English
 5. Not human-fetal sample
 6. Fetal movements not assessed by four-dimensional (4D) sonography
 7. Articles focusing exclusively on fetal facial expressions/movements
 8. Articles not having pertinent titles
-

2.4.1.3 Information sources and data extraction

Literature searches were conducted in PubMed, EMBASE (via OvidSP) and PsycINFO (via EBSCOhost). To scrutinize the literature, a search strategy combining medical subject headings (MeSH) and keywords related to the scope of this review was developed for PubMed (Table S2) and then adapted to the syntax and appropriate subject headings of the other databases. No study design, date or participant demographic restrictions were imposed on the search to ensure literature saturation. After initial searches were completed and duplicate records removed, citations were screened on title and abstract against the eligibility criteria

(Table S1) by two authors (NP, CB), with any discrepancies adjudicated by a third author (DP). Eligible citations were retrieved in full. A standardized data coding system was developed to extract the following information from each study: basic study identifying information, sample size, gestational age, type of movement, duration of US observation, technical parameters of US recording (e.g., frame rate, field of view), additional information about the setting, coding process, data analysis, reported results. Each article was coded by one author (NP) and checked by a separate author (CB). The electronic database search was conducted on March 1st, 2020. An additional hand-search was conducted by one review author (NP) on March 1st, 2020, and then repeated on November 1st, 2020.

Table S2. PubMed search strategy based on MeSH terms and keywords

#1 Fetal Movement MeSH terms
#2 Fetal Development MeSH terms
#3 Fetus MeSH terms
#4 Fetal Title/Abstract
#5 Fetus* Title/Abstract
#6 Foetal Title/Abstract
#7 Foetus* Title/Abstract
#8 Prenatal* Title/Abstract
#9 Antenatal* Title/Abstract
#10 Pregnancy Trimesters MeSH terms
#11 Pregnanc* Title/Abstract
#12 Gestation* Title/Abstract
#13 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12
#14 Movement MeSH terms
#15 Movement* Title/Abstract
#16 Activit* Title/Abstract
#17 Behavio* Title/Abstract
#18 Development* Title/Abstract
#19 Neurodevelopm* Title/Abstract
#20 Neurobehavio* Title/Abstract
#21 Neuro-developm* Title/Abstract
#22 Neuro-behavio* Title/Abstract

#23 #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22

#24 Ultrasonography, Prenatal MeSH terms

#25 Diagnostic imaging Subheading

#26 Ultrason* Title/Abstract

#27 Ultrasou* Title/Abstract

#28 Sonograp* Title/Abstract

#29 Echogr* Title/Abstract

#30 Scan* Title/Abstract

#31 #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30

#32 4D Title/Abstract

#33 4-D Title/Abstract

#34 4-dimension* Title/Abstract

#35 Four-dimension* Title/Abstract

#36 Real-time Title/Abstract

#37 Dynamic Title/Abstract

#38 Kinemat* Title/Abstract

#39 KANET Title/Abstract

#40 #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39

#41 #13 AND #23 AND #31

2.4.2 Results

2.4.2.1 Literature search and study characteristics

Figure S1 demonstrates the selection process used to identify studies included for systematic review. An initial systematic search returned 2833 studies from the three bibliographic databases and 8 from additional sources. After removal of 783 duplicates, 2058 articles were screened for inclusion. Following screening of title and abstract, 244 articles were reviewed as full text. A final number of 56 studies were included in the systematic review. Basic defining features of all studies are summarized in Table S3.

We aggregated studies based on outcome measure:

- Studies reporting on the frequency of fetal movements (n = 30)
- Studies assessing the quality of fetal movements (n = 22)
- Studies assessing the kinematics of fetal movements (n = 4)

In the following section, we provide a narrative synthesis of the methods and results of the included studies, structured by to technical parameters of US recording, duration of US observation, coding system, and reported outcomes.

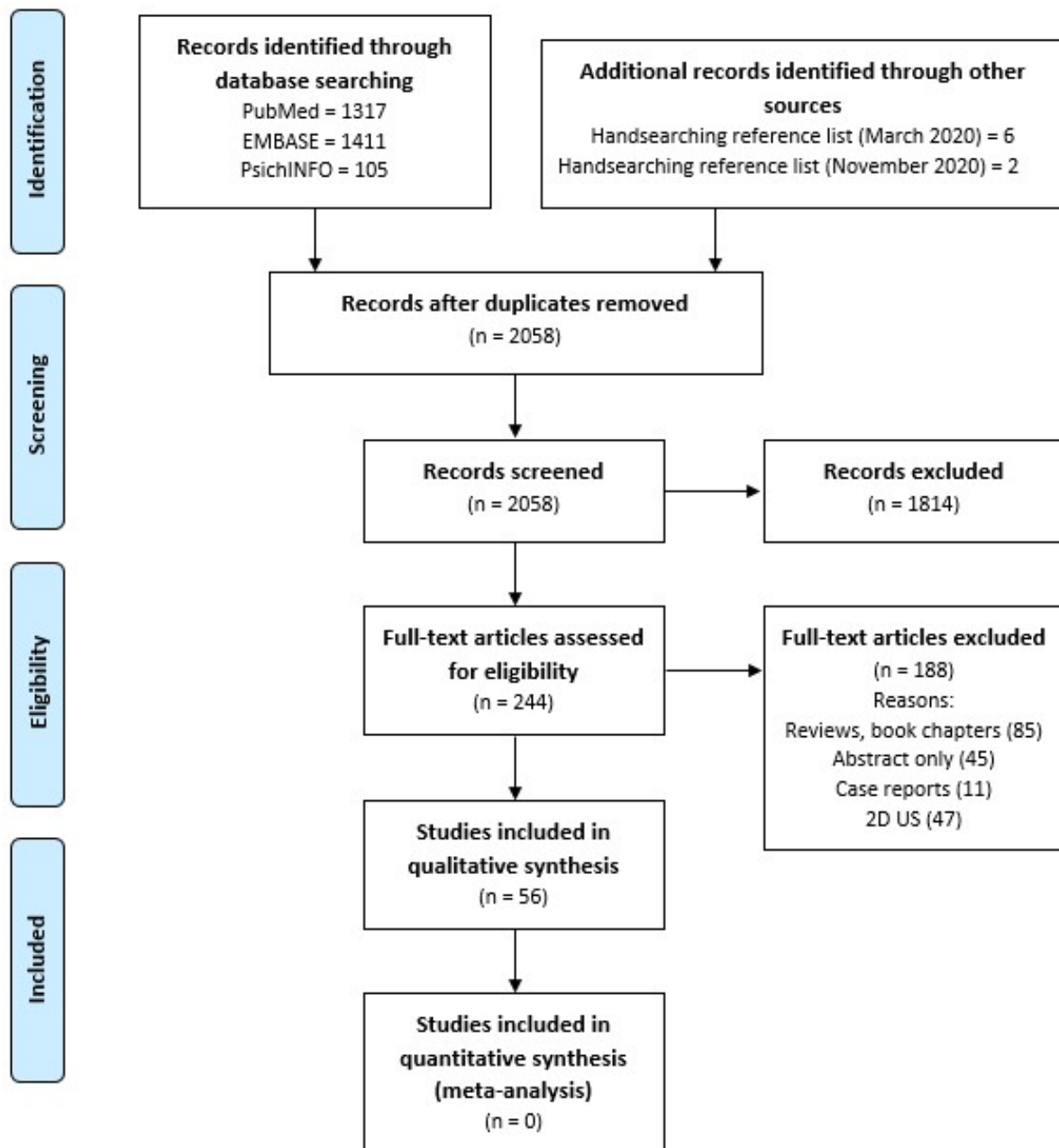


Figure S1. PRISMA flowchart. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart detailing the results of the literature search, article screening and study selection process.

2.4.2.2 Narrative synthesis of studies reporting on the frequency of fetal movements

Thirty studies report on the frequency of fetal movement patterns across trimesters in singletons and twin fetuses.

Studies reporting on the frequency of fetal movements in singletons. Seven studies assessed the spontaneous frequency of fetal patterns during the first trimester in singletons (Andonotopo et al., 2005; Hata et al., 2010; Kurjak et al., 2006; Kurjak et al., 2005; Kurjak et al., 2002; Lebit & Vladareanu, 2011; Yigiter & Kavak, 2006). While these studies agree that movement patterns increase in frequency during the first trimester, they disagree on the relative incidence of specific movement patterns, with some studies showing higher incidence of general movements (Kurjak et al., 2005), and other studies showing higher incidence with isolated arm movements at 10-11 weeks of gestation (Kurjak et al., 2010). Studies assessing the relative frequency of spontaneous movement patterns in second and third trimester also produced mixed results (AboEllail et al., 2018; Andonotopo & Kurjak, 2006; Kim et al., 2010; Kuno et al., 2001; Kurjak et al., 2004; Sajapala et al., 2017). Some studies (AboEllail et al., 2018; Kuno et al., 2001; Sajapala et al., 2017) ranked isolated arm movements as the most frequent pattern, whilst Andonotopo and Kurjak (Andonotopo & Kurjak, 2006) and Kim et al. (Kim et al., 2010) ranked head retroflexion and mouth opening movement as the most common pattern, respectively. Six studies assessed the developmental pattern of hand-to-face movement over the course of the second trimester (Kurjak et al., 2006; Kurjak et al., 2003; Kurjak et al., 2005; Lebit & Vladareanu, 2011; Mori et al., 2018; Yigiter & Kavak, 2006). With one exception (Mori et al., 2018), these studies found a decrease in frequency of hand-to-face movement during the second trimester (Kurjak et al., 2006; Kurjak et al., 2003; Kurjak et al., 2005; Lebit & Vladareanu, 2011; Yigiter & Kavak, 2006). Two studies reported a positive relationship between the frequency of hand-to-face movements, maternal stress (Reissland, Aydin, et al., 2015) and smoking habits (Reissland, Francis, et al., 2015).

The above studies focused on the frequency of hand-to-face movements. Two studies extended the frequency approach to examine the co-occurrence of hand-to-face movements and mouth movements in the second and third trimester (Myowa-Yamakoshi & Takeshita, 2006; Reissland et al., 2014). The logic of these studies was to examine whether mouth opening anticipates the second hand–mouth contact event or is a reaction to it. Both studies found anticipatory mouth movements (Myowa-Yamakoshi & Takeshita, 2006; Reissland et al., 2014). In Reissland et al. (Reissland et al., 2014) but not in Myowa-Yamakoshi et al. (Myowa-Yamakoshi & Takeshita, 2006), the frequency of anticipatory mouth movements increased between the second and the third trimester. Finally, 5 studies assessed how the frequency of fetal movements in the second and third trimester changes in response to external stimuli (Marx & Nagy, 2015, 2017; Ogo et al., 2019; Reid et al., 2017; Reissland et al., 2020). Marx and Nagy (Marx & Nagy, 2015) found that fetuses increase head, arm and mouth movements in response to the touch of the maternal abdomen and decrease arm and head movements to maternal voice. Marx and Nagy (Marx & Nagy, 2017) expanded this approach to different types of abdominal touch, showing differential fetal responses to mother’s touch, father’s touch and stranger’s touch. Using vibroacoustic stimulation (VAS), Ogo et al. (Ogo et al., 2019) reported an increase in blinking and startle movements at 36-40 weeks of gestation, but not earlier. Finally, Reid et al. (Reid et al., 2017) and Reissland et al. (Reissland et al., 2020) reported the frequency of fetal head turns to be higher in the direction of face-like light stimuli compared to non-face-like control light stimuli.

Studies reporting on the frequency of fetal movements in twins. Eight studies assessed the spontaneous frequency of fetal patterns during the first trimester in twins (AboEllail et al., 2018; Degani et al., 2009; Hata et al., 2011; Hata et al., 2012; Kurjak, Talic, Stanojevic, et al., 2013; Mori et al., 2018; Prats et al., 2012; Sasaki et al., 2010). Two studies focused on the frequency of individual movement patterns (AboEllail et al., 2018; Mori et al., 2018), ranking general movements as the most frequent pattern at 12-13 weeks of gestation. Other studies evaluated the relative frequency of inter-twin contacts. Three studies compared the frequency of inter-twin contacts between monochorionic diamniotic (MD) and dichorionic diamniotic (DD) twins in the first trimester of pregnancy (Hata et al., 2011; Hata et al., 2012; Sasaki et al.,

2010). Assessing fetuses at 10-11 weeks of gestation, Sasaki et al. (Sasaki et al., 2010) reported more frequent inter-twin contacts in MD compared to DD twins, whilst Hata et al. (Hata et al., 2012) reported no significant differences between MD and DD. Results at 12-13 weeks of gestation were also inconsistent. Hata et al. (Hata et al., 2012) reported head to arm contacts (but no other inter-twin contacts) to be higher in MD compared to DD twins. In contrast, Sasaki et al. (Sasaki et al., 2010) and Hata et al. (Hata et al., 2011) found no significant differences between MD and DD twins. Two studies (Kurjak, Talic, Stanojevic, et al., 2013; Prats et al., 2012) mapped inter-twin contacts as a function of complexity, from twins touching each other without evidence of reaction, to movement of the co-twin with one second of touch by the co-twin (first reaction), to complex contacts lasting more than 5 sec and involving repetitive action and reaction of the co-twins. Both studies found the complexity of intertwin contacts to increase starting from 12 weeks of gestation. Finally, Degani et al. (Degani et al., 2009) examined the relationship between antenatal inter-twin activity and the post-natal infant twins' temperament. Prenatal inter-twin differences in total number of movements were reported to be in good correlation with postnatal maternal reports on infants' temperament.

2.4.2.3 Narrative synthesis of studies assessing the quality of fetal movements

Twenty-two papers assessed the quality of fetal movements using the Kurjak Antenatal Neurodevelopmental Test (KANET) (Kurjak et al., 2010). Seventeen studies reported significant differences in KANET scores between low-risk and high-risk fetuses in the third trimester of pregnancy (Abo-Yaqoub et al., 2012; Antsaklis et al., 2017; Athanasiadis et al., 2013; Dieb et al., 2019; Hata, Kanenishi, Mori, AboEllail, Hanaoka, Koyano, Kato, & Kusaka, 2019; Hata, Kanenishi, Mori, AboEllail, Hanaoka, Koyano, Kato, Kusaka, et al., 2019; Honemeyer et al., 2013; Kurjak et al., 2010; Kurjak et al., 2008; Kurjak, Talic, Honemeyer, et al., 2013; Miskovic et al., 2010; Neto, 2015; Predojević et al., 2014; Talic et al., 2011; Talic et al., 2012; Tinjic, 2019; Vladareanu et al., 2012). The remaining 5 studies reported differences in KANET scores for specific movement types in low-risk fetuses, specifically Japanese versus

Croatian fetuses (Hanaoka et al., 2016), male versus female fetuses (Hata et al., 2016a), fetuses of primiparas versus multiparas (Hata et al., 2016b), and dichorionic twin pregnancies versus monochorionic twin pregnancies (Bot et al., 2020). Hata et al. (Hata, Kanenishi, Mori, AboEllail, Hanaoka, Koyano, Kato, Kusaka, et al., 2019) assessed the usefulness of a simplified version of the KANET test (Mini KANET) for the prediction of postnatal developmental disabilities.

2.4.2.4 Narrative synthesis of studies assessing the kinematics of fetal movements

Four papers (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007) assessed the kinematics of fetal arm movements to determine whether early in the second trimester of pregnancy (14-22 weeks) fetuses prospectively calibrate their reaching patterning to property of the target (mouth region, eye region, co-twin, no target) (Becchio et al., 2018; Rosenbaum et al., 2012). Examinations were conducted over a period of 20 minutes with the ultrasound probe set to 4 frames per second capturing the torso, arm and head of the fetus. Extracted frames were manually digitized and wrist displacement was computed with reference to the fetus body axes, using the intra-ocular distance as measurement unit. Zoia et al. (Zoia et al., 2013) specify that a target was considered as to be reached each time the hand stopped on the target for at least two consecutive frames. Three studies applying this method in singletons found differences in the calibration of movements towards the mouth and the eye region at 22 weeks of gestation, but not at 14 and 18 weeks (Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007). One study (Castiello et al., 2010) expanded this approach to twins, reporting calibration effects already at 14 and 18 weeks of gestations.

Table S3. Studies assessing fetal movements based on outcome measure

Authors (years)	Fetuses	GA weeks	Sample	Type of fetal movements	Frames/s hz	Duration min	Coders n.	Intercoder reliability
Studies reporting on the frequency of fetal movements								
AboEllail, Kanenishi, Mori, Noguchi, Marumo, Hata (2018)	Singleton Twin	12-19	58 singletons 25 twins (12MD; 13DD)	Head anteflexion, head retroflexion, body rotation, hand-to-face, general movements, arm movements, leg movements, mouth movements	NS	15	1	NA
Andonotopo, Medic, Salihagic-Kadic, Milenkovic, Maiz, Scazzocchio (2005)	Singleton	6-14	50	General movements, startle, isolated arm movements, isolated leg movements, stretching, hand-to-face, head retroflexion, hand anteflexion, head rotation	0.5	15	NS	NS
Andonotopo, Kurjak (2006)	Singleton	30-40	100 (50 IUGR; 50 AGA)	Isolated eye blinking, mouthing, yawning, tongue expulsion, grimacing, swallowing, smiling, sucking, head retroflexion, head rotation, head anteflexion, hand-to-head, hand-to-eye, hand-to-mouth, hand-to-face, hand-to-ear	0.5	30	NS	NS
Degani, Leibovitz, Shapiro, Ohel (2009)	Twin	11-14	26 pairs (MD, DD: NS)	Isolated arm movements, isolated leg movements, stretching, head movements, general body movements	NS	20	1	NA
Hata, Kanenishi, Sasaki (2010)	Singleton	10-11, 12-13	17	Isolated arm movements, isolated leg movements, short trunk (trunk movement for <20 sec), long trunk (trunk movement for >20 sec), jumping movements	2.6-4	10	1	NA
Hata, Kanenishi, Sasaki, Yanagihara (2011)	Twin	12-13	14 pairs (6 MD; 8DD)	No reaction movements (twins appear to touch each other but there is no clear reaction), reaction movements (twins appear to touch each other and there is a clear reaction by the co-twin)	2.6-4	30	1	NA
Hata, Sasaki, Yanagihara (2012)	Twin	10-11, 12-13	14 pairs (6 MD; 8DD)	Inter-twin contact (head-to-head, head-to-arm, head-to-trunk, head-to-leg, arm-to-arm, arm-to-trunk, arm-to-leg, trunk-to-trunk, trunk-to-leg, leg-to-leg)	2.6-4	30	1	NA
Kim, Lee, Chung, Lee, Lee, Choi, Lee (2010)	Singleton	28-32	40	Sucking, swallowing, isolated limb movement	NS	30	NS	NS
Kuno, Akiyama, Yamashiro, Tanaka, Yanagihara, Hata (2001)	Singleton	14-18	11	Head (retroflexion, rotation, and anteflexion), mouth (jaw opening and yawning), arm (flexion and extension), trunk (stretching and bending or flexion and extension), and leg (flexion and extension) movements	0.5	60	NS	NS

Kurjak, Vecek, Hafner, Bozek, Funduk-Kurjak, Biserka, Ujevic (2002)	Singleton	6-12	98	Gross body movements, limb movements, complex limb movements	NS	Up to 2	NS	NS
Kurjak, Azumendi, Veček, Kupešic, Solak, Varga, Chervenak (2003)	Singleton	12-16, 30-33	25	Isolated hand movement, subtypes of hand movements (hand-to-head, hand-to-mouth, hand-near-mouth, hand-to-face, hand-near-face, hand-to-eye, hand-to-ear), facial expressions (isolated blinking, mouthing movements, mouth and eyelid movement, yawning, tongue expulsion, pouting, smiling, scowling)	0.5	15	NS	NS
Kurjak, Stanojevic, Andonotopo, Salihagic-Kadic, Carrera, Azumendi (2004)	Singleton	33-35	10	Facial expression, hand movements (hand-to-head, hand-to-mouth, hand-to-eye, hand-to-ear)	0.5	15	NS	NS
Kurjak, Stanojevic, Andonotopo, Scazzocchio-Duenas, Azumendi, Carrera (2005)	Singleton	8-35	100	Sideway bending, general movements, startle, isolated arm movements, isolated leg movements, hiccup, suckling, swallowing, hand to face contact, head retroflexion, head antelexion, head rotation, breathing movement, stretching, jaw opening, yawning	0.5	30	NS	NS
Kurjak, Andonotopo, Hafner, Kadic, Stanojevic, Azumendi, Ahmed, Carrera, Troyano (2006)	Singleton	7-40	100	1st trimester: general movements, startle, stretching, isolated arm movements, isolated leg movements, head retroflexion, head rotation, head antelexion; 2nd-3rd trimester: isolated eye blinking, mouthing, yawning, tongue expulsion, grimacing, swallowing, head retroflexion, head rotation, head antelexion, hand-to-head, hand-to-eye, hand-to-mouth, hand-to-face, hand-to-ear	0.5	30	NS	NS
Kurjak, Talic, Stanojevic, Honemeyer, Serra, Prats, Di Renzo (2013)	Twin	8-16	20 pairs (MD, DD: NS)	First reach and touch, first reaction, complex contact, mouth contact, diffuse interactions	NS	Up to 20	2	Coefficient of variation ranging from 91% to 100%
Lebit, Vladareanu (2011)	Singleton	7-38	144	1st trimester: general movements, startle, stretching, isolated arm movements, isolated leg movements, head retroflexion, head rotation, head antelexion; 2nd-3rd trimester: facial expression, isolated eye blinking, mouthing, yawning, tongue expulsion, grimacing, swallowing, complex movements	NS	15-20	NS	NS
Marx, Nagy (2015)	Singleton	21 -33	23	Arms movements, head movements, mouth movements, arms crossing, hand-to-body, yawning	4	27	2 (one blind)	k = .77

Marx, Nagy (2017)	Singleton	20-35	28	Arm movements, head movements, hands touching the body/face, arms crossed, hands crossed, body rotation, mouth movements, yawning, tongue protrusion, sucking, hiccups, breathing, stretch, kicking of the legs	NS	24	2	Frequency: $k = .91$; Duration: $k = .92$
Mori, Kanenishi, Aboellail, Nitta, Noguchi, Marumo, Hata (2018)	Singleton Twin	12-19	58 singletons 25 pairs (12MD; 13DD)	Head anteflexion, head retroflexion, body rotation, hand-to-face, general movements, arm movements, leg movements, mouth movements	NS	15	1	NA
Myowa-Yamakoshi, Takeshita (2006)	Singleton	19-35	27	Hand-to-face, mouth movements (opening, closing)	2-4	20	2 (one blind)	$k = .78$
Ogo, Kanenishi, Mori, Aboellail, Hata (2019)	Singleton	24-40	68	Mouthing, yawning, tongue expulsion, smiling, scowling, blinking, back arch, jerky arm movements, hand-to-face, startle	NS	3 pre + 3 post VAS	NS	NS
Prats, Serra, Fournier, Baulies, Andonotopo, Kurjak (2012)	Twin	8-16	20 pairs (MD, DD: NS)	First reach and touch, first reaction, complex contact, mouth contact, diffuse interactions	NS	20	2	Coefficient of variation ranging from 91% to 100%
Reid, Dunn, Young, Amu, Donovan, Reissland (2017)	Singleton	33-36	39	Head turns	NS	5 sec/trial	2 (one blind)	$k = .79$
Reissland, Francis, Aydin, Mason, Schaal (2014)	Singleton	24-36	15	Hand-to-face (upper face, side face, lower face, mouth area), mouth movements (opening, closing)	NS	20	2	Touch behaviors: $k = .87$; mouth movements: $k = .86$
Reissland, Aydin, Francis, Exley (2015)	Singleton	24-36	15	Hand-to-head, hand-to-face (with left and right arms)	NS	10	2 (one blind)	$k = .96$
Reissland, Francis, Kumarendran, Mason (2015)	Singleton	24-36	20 (4 smokers; 16 non-smokers)	Hand-to-face, mouth movements	2	15-20	2	Nonsmokers: $k = .90$ (mouth), $k = .88$ (facial); Smokers: $k = .87$ (mouth); $k = .90$ (facial)
Reissland, Wood, Einbeck, Lane (2020)	Singleton	33	9	Head turns	NS	NS	2 (one blind)	$k = .87$
Sajapala, Aboellail, Kanenishi, Mori, Marumo, Hata (2017)	Singleton	14-16, 17-19	29	Head anteflexion, head retroflexion, body rotation, hand to face movements, general movements, isolated arm movements, isolated leg movements, mouth movements	NS	15	2	Inter-class correlation coefficients $> .88$

Sasaki, Yanagihara, Toshihiro, Naitoh, Hata (2010)	Twin	10-11, 12-13	14 pairs (MD, DD: NS)	Inter-twin contact: head-to-head, head-to-arm, head-to-trunk, head-to-leg, arm-to-arm, arm-to-trunk, arm-to-leg, trunk-to-trunk, trunk-to-leg, leg-to-leg	2.6-4	30	1	NA
Yigiter, Kavak (2006)	Singleton	11-40	63	1st trimester: general movements, startle, stretching, isolated arm movements, isolated leg movements, head retroflexion, head rotation, head anteflexion; 2nd-3rd trimester: facial expression, isolated eye blinking, yawning, tongue expulsion, grimacing, swallowing, smiling, sucking, head retroflexion, head rotation, head anteflexion, hand-to-head, hand-to-eye, hand-to-mouth, hand-to-face, hand-to-ear	0.5	30	NS	NS

Studies reporting the quality of fetal movements

Abo-Yaqoub, Kurjak, Mohammed, Shadad, Abdel-Maaboud (2012)	Singleton	20-38	80 (40 HR ¹ ; 40 LR)	Movements included in KANET test *	NS	NS	NS	NS
Antsaklis, Porovic, Daskalakis, Kurjak (2017)	Singleton	28-38	80 (40 diabetic mothers; 40 non-diabetic mothers)	Movements included in KANET test *	NS	15-20	NS	NS
Athanasiadis, Mikos, Tambakoudis, Theodoridis, Papastergiou, Assimakopoulos, Tarlatzis (2013)	Singleton	HR: 31.8 (mean); LR: 30.3 (mean)	152 (74 HR ² ; 78 LR)	Movements included in KANET test *	NS	20-30	NS	NS
Bot, Vladareanu, Burnei, Munteanu, Calo, Vladareanu (2020)	Twin	28, 30-32, 34-36	91 (24 MC; 67 DC)	Movements included in KANET test *	NS	15-20	NS	NS
Dieb, Salam, Shaheen, Shaeer (2019)	Singleton	28-38	50 (25 euthyroid mothers; 25 hypothyroid mothers)	Movements included in KANET test *	NS	15-20	NS	NS
Hanaoka, Hata, Kanenishi, AboEllail, Uematsu, Konishi, Kusaka, Noguchi, Marumo, Vasilj, Kurjak (2016)	Singleton	28-38	167 (89 Japanese; 78 Croatians)	Movements included in KANET test *	NS	15-20	NS	NS

Hata, Hanaoka, AboEllail, Uematsu, Noguchi, Kusaka, Kurjak (2016a)	Singleton	28-39	112 (53 females; 59 males)	Movements included in KANET test *	NS	15-20	NS	NS
Hata, Hanaoka, AboEllail, Uematsu, Noguchi, Kusaka, Kurjak (2016b)	Singleton	32-36	105 (44 primiparas; 61 multiparas)	Movements included in KANET test *	NS	15-20	NS	NS
Hata, Kanenishi, Mori, AboEllail, Hanaoka, Koyano, Kato, Kusaka (2019)	Singleton	28-38	353	Movements included in KANET test *	NS	15-20	NS	NS
Hata, Kanenishi, AboEllail, Mori, Koyano, Kato, Kusaka (2019)	Singleton	28-38	40 (30 healthy; 10 psychotropic-drug-administered mothers)	Movements included in KANET test *	NS	15-20	NS	NS
Hata, Kanenishi, Mori, AboEllail, Hanaoka, Koyano, Kato, Kusaka, Kurjak (2019)	Singleton	28-38	353	Movements included in KANET test *	1-3	15-20	NS	NS
Honemeyer, Talic, Therwat, Paulose, Patidar (2013)	Singleton	28-38	56 (32 HR ³ ; 24 LR)	Movements included in KANET test *	NS	Up to 30	NS	NS
Kurjak, Miskovic, Stanojevic, Amiel-Tison, Ahmed, Azumendi, Vasilj, Andonotopo, Turudic, Salihagic-Kadic (2008)	Singleton	all trimesters (weeks NS)	100 LR (1 st study) 120 HR ³ (2 nd study)	Movements included in KANET test *	NS	30	NS	NS
Kurjak, Abo-Yaqoub, Stanojevic, Yigiter, Vasilj, Lebit, Shaddad, Ahmed, Kavak, Miskovic, Vladareanu, Barisic, Azumendi, Younis, Pooh, Kadic (2010)	Singleton	20-38	288	Movements included in KANET test *	NS	NS	NS	NS
Kurjak, Talic, Honemeyer, Stanojevic, Zalud (2013)	Singleton	24-38	869 (596 HR ³ ; 273 LR)	Movements included in KANET test *	NS	Up to 20	NS	NS
Miskovic, Vasilj, Stanojevic, Ivanković, Kerner, Tikvica (2010)	Singleton	20-36	226 (116 HR ³ ; 110 LR)	Movements included in KANET test *	NS	NS	NS	NS

Neto (2015)	Singleton	28-38	51 (17 HR ³ ; 34 LR)	Movements included in KANET test *	NS	Up to 30	NS	NS
Predojevic, Talić, Stanojević, Kurjak, Salihagić Kadić (2014)	Singleton	31, 32, 37-39	5 IUGR	Movements included in KANET test *	NS	30	NS	NS
Talic, Kurjak, Ahmed, Stanojevic, Predojevic, Salihagic Kadic, Di Renzo (2011)	Singleton	26-38	620 (520 HR ³ ; 100 LR)	Movements included in KANET test *	NS	Up to 30	NS	NS
Talic, Kurjak, Stanojevic, Honemeyer, Badreldeen, DiRenzo (2012)	Singleton	32-36	240 (140 VM; 100 LR)	Movements included in KANET test *	NS	Up to 30	NS	NS
Tinjić (2019)	Singleton	28-38	141 (12 HR ³ ; 129 LR)	Movements included in KANET test *	NS	Up to 30	NS	NS
Vladareanu, Lebit, Constantinescu (2012)	Singleton	24-38	196 (135 HR ³ ; 61 LR)	Movements included in KANET test *	NS	15-30	NS	NS

Studies reporting the kinematics of fetal movements

Castiello, Becchio, Zoia, Nelini, Sartori, Blason, D'Ottavio, Bulgheroni, Gallese (2010)	Twin	14, 18	5 pairs (MD, DD: NS)	Hand-to-mouth, hand-to-eye, movements to the uterine wall, hand-to-back of the co-twin, hand-to-head of the co-twin	4	20	2 (both blind)	k = .93
Parma, Brasselet, Zoia, Bulgheroni, Castiello (2017)	Singleton	14, 18, 22	29	Hand-to-eye, hand-to-mouth, movements to the uterine wall	4	20	2	NS
Zoia, Blason, D'Ottavio, Bulgheroni, Pezzetta, Scabar, Castiello (2007)	Singleton	14, 18, 22	8	Hand-to-mouth movements, hand-to-eye movements	4	20	2	NS
Zoia, Blason, D'Ottavio, Biancotto, Bulgheroni, Castiello (2013)	Singleton	14, 18, 22	8	Hand-to-mouth movements, hand-to-eye movements	4	20	3	Intra-class correlation coefficients > .99

AGA: appropriate-for-gestational age; DD/DC: dichorionic diamniotic twins; HR: high-risk; IUGR: intrauterine growth restriction; k: Cohen's Kappa; LR: low-risk; MD/MC: monochorionic diamniotic twins; NA: not applicable; NS: not specified; VM: ventriculomegaly.

* Movements included in KANET test: isolated head antelexion, cranial sutures and head circumference, isolated eye blinking, facial alteration and mouth opening, isolated leg movement, isolated hand movement, hand-to-face movements, finger movements and thumb position, and gestalt perception of general movements.

¹ The risk reported is for "neurological abnormalities". Inclusion criteria in the high-risk group are not further specified.

² Reported criteria to define the high-risk group are: IUGR; women with gestational diabetes or type I diabetes; women with preexisting hypertension or hypertension developed during pregnancy.

³ In these studies the high-risk group is reported as a highly heterogeneous group including a variety of pre-existing maternal conditions (e.g., hyperthyroidism, cyanotic heart disease, LES), and pregnancy associated conditions (e.g., preeclampsia, oligohydramnios, preterm labor and/or premature preterm rupture of membranes).

Chapter 3 - Towards the Use of 2D Video-Based Markerless Motion Capture to Track Fetal Movements

3.1 Introduction

Accurate quantification of behavior is essential to understanding the brain (Krakauer et al., 2017; Pereira et al., 2020), and fetal behavior is likely to directly reflect the development and maturation of the fetal brain and central nervous system. The introduction of four-dimensional ultrasonography (4D US) has enabled recording of dynamic movement in fetuses by providing continuous visualization of fetal movements. However, as discussed in Chapter 2, progress in characterizing fetal movements has been hindered by the lack of tools to objectively and quantitatively assess fetal behavior.

In motor control studies, the gold standard for reliable movement analysis is the fitting of reflective markers to assist with computer-based motion tracking. Unfortunately, this approach is not applicable to fetal movements. Previous studies reporting on the kinematics of fetal movements extracted motion parameters from manual digitization of 3D surface rendered images (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007). However, annotating 4D US scans manually is time-consuming, labor-intensive, and error-prone. In Chapter 2, we identified the application of deep neural networks as a promising direction to improve the quantitative analysis of fetal movements. Markerless pose estimation of user-defined body parts, using deep learning architectures, has proven an effective tool for measuring behavior (Mathis et al., 2018; Mathis et al., 2020). Pose estimation refers to methods for measuring posture, while posture denotes the geometrical configuration of body parts. A pose estimation algorithm can be viewed as a function that maps frames from a video into the coordinates of body parts. By decomposing objects into key points with semantic meaning - such as body parts in videos of human subjects - a high-dimensional video signal can be converted into a collection of time series describing the

movement of each key point. Compared to raw video, this representation is easy to analyze and semantically meaningful for investigating behavior (Mathis et al., 2020). The dataset defines the input-output relationships that the model should learn: in pose estimation, a particular pose (output) should be predicted for a particular image (input). The model's parameters (weights) are iteratively updated by the optimizer to minimize the loss function (Mathis et al., 2020). Thereby the loss function measures the quality of a predicted pose (in comparison to ground truth data if existing). The only real requirement is the ability of the human to be able to reliably label key points.

To be useful, pose estimation tools need to be as good as human annotation of frames. That means the algorithm performance should be assessed (i.e. benchmarked) by extracting performance measures and then comparing with the human performance. Methods for 2D markerless motion capture are usually benchmarked on large-scale labeled datasets (Dang et al., 2019), which may render deep learning approaches infeasible for diverse pose estimation tasks. To provide a relatively easy-to-use toolbox for motion tracking - aiming at benchmarking on smaller datasets - a few years ago Mathis and colleagues introduced DeepLabCut. The structure and training of the network have been described in detail elsewhere (Mathis et al., 2018; Nath et al., 2019). Briefly, DeepLabCut is a deep neural network pre-trained on ImageNet, a massive open database of images used for image recognition research (Deng et al., 2009). This allows to reduce the amount of data needed for training and to achieve close-to-human performance in tracking movements with only a few hundred labeled training images. DeepLabCut detects multiple body parts at the same time by using the location of each body part to assist in the localization of the others. As a supervised learning algorithm, DeepLabCut starts with a labeled dataset, then exploits this information to estimate how probable each image region is to be a given body part. In other words, the network predicts the probability that a labeled part, i.e. a fetus' body part, is in a particular pixel. In recent years, DeepLabCut has been tested and calibrated across species (e.g. rats, flies, horses, and infants) and across activities (e.g. locomotion and hand reaching), however, it has never been used to extract key body part locations from human fetuses.

This work aimed to expand DeepLabCut for fetus pose estimation and generate the first annotated dataset of fetal movements. As a first step, we collected a dataset of 4D US data consisting of 2000 frames collected over the 1st and 2nd trimesters of pregnancy. To take into account the potential human variability in manually annotating fetal body parts, we defined a consensus annotation based on the annotations of three independent human labelers over 2000 frames. Next, we trained our network on consensus annotation and compared its performance to that of individual human annotators. The annotated dataset is released to serve as a benchmark for future algorithm development. To additionally benchmark our network, we then compared its performance against that of the fourth human annotator whose annotations were not used to build the consensus annotation. Finally, we demonstrated the potential of our network for kinematic analysis of fetal movement by using it to compute wrist velocity.

3.2 Methods

3.2.1 Data resource

3.2.1.1 Study population

The research protocol was approved by the local ethics committee (ASL3 Genovese) and complied with the principles of the revised Helsinki Declaration (World Medical, 2013). All pregnant women participating received written information and gave their informed written consent to participate in the study.

The 37 women with a singleton pregnancy enrolled in this study were a convenient sample of low-risk pregnant women who attended the Institute of Child Health IRCCS Giannina Gaslini. The designation of “low risk” for fetuses was made by the gynecologist during the initial obstetric visit based on maternal medical history. All the 37 fetuses were examined the first time at 11-14 gestational weeks. Twelve of them were then re-examined at weeks 19-21. Gestational age was calculated based on the last menstruation period.

3.2.1.2 Image data collection

After the routine 2D US examinations (including fetal biometry measurements to demonstrate normal growth and no anomalies), each fetus was recorded with 4D US for 1 to 4 minutes to obtain sufficient behavioral data. All scans were performed by an experienced sonographer. To visualize fetal upper limb movements, the transducer of transabdominal 4D US was kept as still as possible and the region of interest (ROI) was adapted so that the upper body of each fetus was obtained, including head, thorax, and at least one upper limb with the hand. This choice is driven by clinical considerations: monitoring arms movements is of particular interest for evaluating fetal motor development. 4D examinations were performed using a GE Healthcare Voluson Ultrasound System E10 equipped with an eM6CG3 Volumetric Transducer at an average frame rate across scans of 16 Hz in 1st trimester fetuses and 11 Hz in 2nd trimester fetuses. Due to the fetal growth, a lower frame rate had to be set in 2nd trimester to ensure the visualization of both the face and the upper limb. 4D scans were then exported to audio video interleave (AVI) format and saved to an external hard drive for offline analysis.

3.2.1.3 Image data annotation

Before extracting and labeling selected frames from the recorded videos, we cropped the boundary region of the data (in each direction) to remove black borders introduced in the recording process. To optimize GPU memory during training and to speed up the benchmarking process, we then downsampled the frames by a factor of 25% (obtaining 162×162 pixels images).

To create the annotated dataset, we first extracted 2000 frames from recorded videos. Frames were manually selected to reflect the diversity of the behavior with respect to fetal postures, limb positions, and background conditions. Then, 12 fetal anatomical landmarks were identified to be manually labeled. By using the DeepLabCut toolbox (Mathis et al., 2018), the position of the following 12 body parts (key points) was labeled in 2000 frames, when visible: right wrist (wrist_R), left wrist (wrist_L), right elbow

(elbow_R), left elbow (elbow_L), right shoulder (shoulder_R), left shoulder (shoulder_L), right ear (ear_R), left ear (ear_L), right eye (eye_R), left eye (eye_L), nose, and mouth.

We collected manual annotations of the same dataset from 3 independent labelers (a sonographer and two trained users) who were instructed to identify body parts in a principled and consistent way. The body parts were labeled according to the following guidelines: (1) body parts of the limbs (shoulders, elbows, and wrists) should be located at the center of the joint rotation; (2) as a rule of thumb, ear, eye, and nose body parts should be located at the entrance of the ear canal, the center of the eyeball, in the middle position between the entrances of the two nostrils, respectively; however, due to poor quality 4D images, the position of face parts was often challenging; (3) a body part was annotated if its position was visible, whereas it was skipped if it was not visible due to shadowing by other body parts or restriction in the field of view. The human annotation time of the entire dataset was estimated at around 25 hours per labeler.

3.2.1.4 Generation of a *consensus annotation*

The surface rendering quality of 4D ultrasound is dependent on temporal resolution: the higher the sampling rate, the lower the quality of the rendered image. For this reason, the operation of identifying fetal body parts in 4D US images can be challenging for human labelers. To address human variability in manual annotation, the same dataset (i.e., 2000 frames) was thus annotated by three independent labelers. In a subsequent stage, the annotations were combined to create a *consensus annotation*. A *consensus annotation* was assigned when at least two out of three labelers agreed. The final *consensus annotation* dataset was generated by computing the average position of each body part across the two/three human labelers. This procedure was repeated across all the body parts. The *consensus annotation* dataset was then used to quantify the limits of human performance and to benchmark DeepLabCut.

3.2.1.5 Measuring performance against *consensus annotation*

In line with other studies (Giovannucci et al., 2019; Mathis et al., 2018), we compared the performance of DeepLabCut and the performance of each human labeler against the *consensus annotation*. Performance was evaluated using two accuracy metrics: the F1 score and the root mean square error (RMSE).

F1 score. Key points annotations were classified as true positive (TP), true negative (TN), false positive (FP), and false negative (FN) detections relative to the consensus annotations. A key point was defined as a TP if the frame included the corresponding consensus annotation, regardless of its location in the image. A key point was defined as a FP if the frame did not include the corresponding consensus annotation. From TP, FP and FN, we extracted Precision and Recall. Precision measures the fraction of correct detections out of all detections and can be obtained by:

$$Precision = \frac{TP}{TP + FP}$$

Recall weights the fraction of positives that are correctly detected and can be obtained by:

$$Recall = \frac{TP}{TP + FN}$$

The F1 score combines Precision and Recall via their geometric mean:

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

RMSE. For TP, we further evaluated the accuracy of the annotation as the RMSE between the annotation and the corresponding consensual annotation. The distance error of position estimation (in pixels) was calculated separately for each key point.

3.2.1.6 Algorithm and human performances evaluation against *consensus annotation*

DeepLabCut performances against *consensus annotation* were evaluated in terms of detection F1 score and RMSE by performing a 10-fold cross-validation analysis. We first randomly partitioned the dataset (i.e. the *consensus annotation* dataset) into 10 equal size folds. One of the ten folds is then removed from the training set and used as validation set (i.e. for testing the model), whereas the remaining 9 folds are used as training set. This cross-validation process is repeated 10 times, with each of the 10 folds used as validation. For each key point, averaged values of the algorithm detection accuracy (i.e. F1 score, Table 3.1) and the RMSE (Table 3.2) across 10 validation sets were computed.

We then measure human performances across the 10 validation sets to compare DeepLabCut performances with the human ones. For each labeler and each key point, we computed the averaged values of the F1 score (Table 3.1) and the RMSE (Table 3.2) across 10 validation sets.

3.3 Results

3.3.1 Comparing DeepLabCut against manual annotations

F1 score. The detection accuracies of both the algorithm and human labelers were investigated by judging whether a label exists in the frame or not, with consensus annotations as a reference. This was quantified with the F1 score. Results of the three labelers (i.e. L1, L2, and L3) and DeepLabCut (DLC) are shown in Figure 3.2, for each key point. The range of human performance (blue, orange, and green bars in Figure 3.2) was 0.90-0.93. All three labelers performed similarly on average (0.91 ± 0.05 , 0.90 ± 0.04 and 0.93 ± 0.02 , mean $\pm$ SD). Detection scores of shoulders reveal a slightly higher variability across labelers, compared to other body parts. DeepLabCut (red bars in Figure 3.2) had an average performance of 0.91 ± 0.02 (mean $\pm$ SD). Performance was similar across all fetal body parts with slightly lower performances in the detection of left shoulder (F1 score mean $\pm$ SD = 0.88 ± 0.03 px), right shoulder (F1 score mean $\pm$

SD = 0.88 ± 0.03 px), left ear (F1 score mean $\pm$ SD = 0.89 ± 0.02 px), and right ear (F1 score mean $\pm$ SD = 0.89 ± 0.02 px).

RMSE. For TP, we further computed RMSE of position estimation by human labelers and DeepLabCut against consensus annotation. The RMSE of position estimation by human labelers (blue, orange, and green bars in Figure 3.3) was in the range of 2.74 – 2.84 pixels, with a mean RMSE of 2.74 ± 0.32 (mean $\pm$ SD), 2.84 ± 0.32 (mean $\pm$ SD), and 2.78 ± 0.28 (mean $\pm$ SD) pixels for labeler 1, 2 and 3. The RMSE of error of position estimation by DeepLabCut (red bars in Figure 3.3) was 3.17 ± 0.20 pixels (mean $\pm$ SD). Distances between DeepLabCut and consensual annotations were similar across all key points, with slightly higher error values for the left ear (RMSE mean $\pm$ SD = 3.49 ± 0.42 px) and right ear (RMSE mean $\pm$ SD = 3.39 ± 0.26 px). The mean delta value between all labelers and DeepLabCut was 0.38 pixels. For reference, such a size corresponds to $\sim 0.23\%$ of the image height (i.e. 162 x 162 pixels).

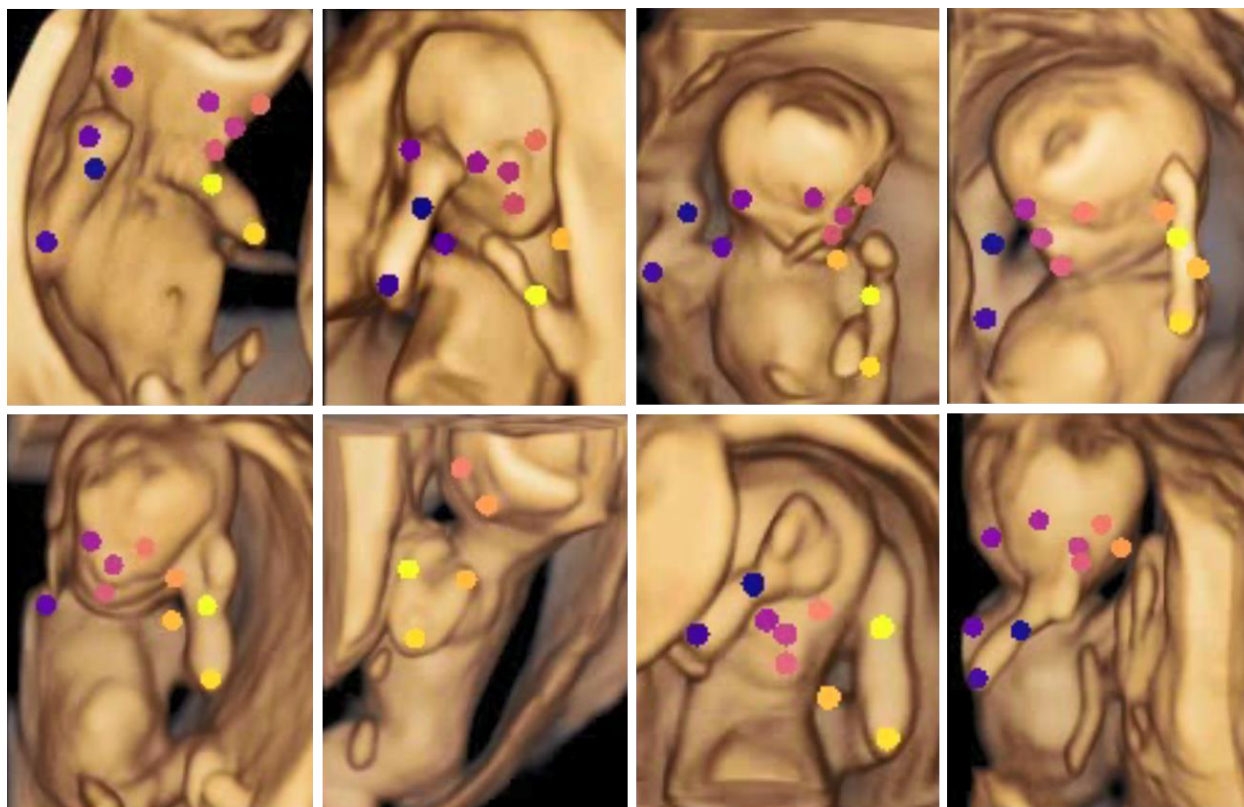


Figure 3.1 Examples of frames and labeled body parts predicted by DeepLabCut. Note that for display purposes, the frames are cropped with respect to their original size.

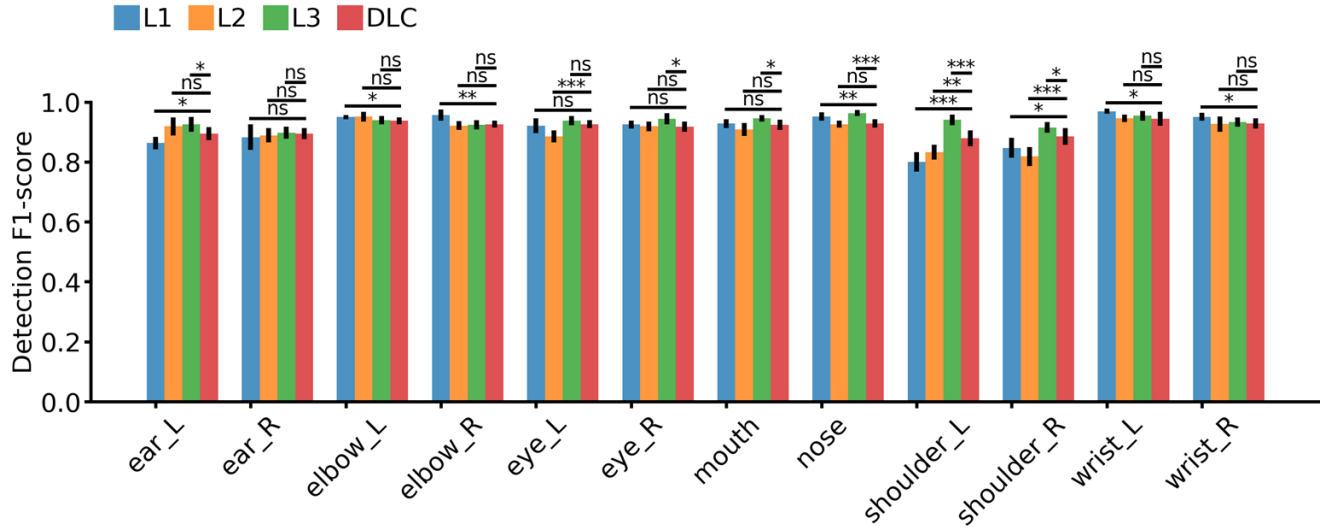


Figure 3.2 Mean F1 scores of predicted (red bars) and human labelers' (blue, orange, and green bars) detections of each key point computed against the consensus annotations. Error bars indicate standard deviation (SD). Asterisks show two-sided independent samples t-test significance, Holm-Bonferroni corrected. * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, *ns* indicates $p > 0.05$.

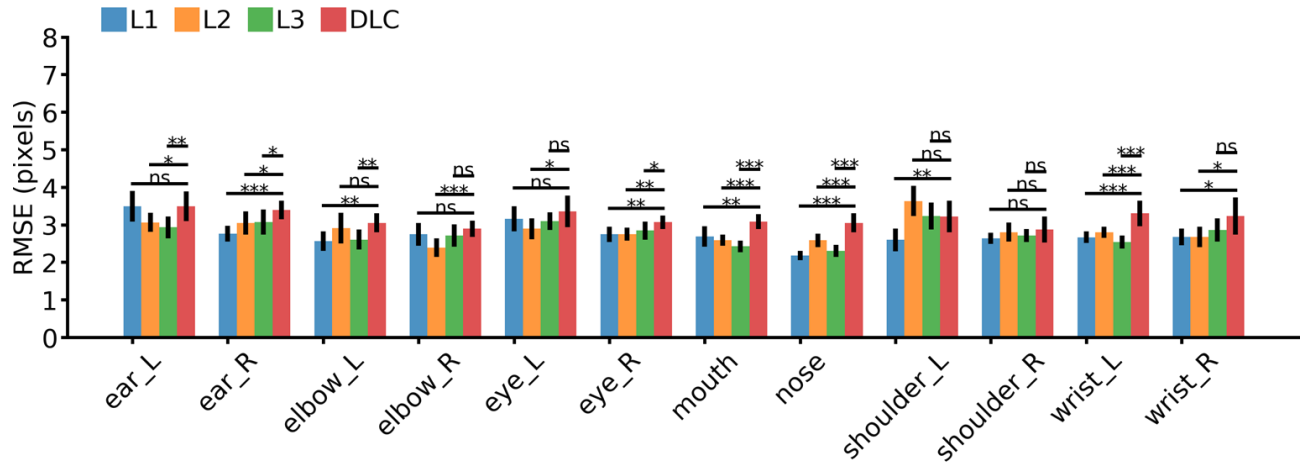


Figure 3.3 Averaged root mean square error (RMSE) of predicted (red bars) and human labelers' (blue, orange, and green bars) positions of each key point computed against the consensus annotations. Error bars indicate standard deviation (SD). Asterisks show two-sided independent samples t-test significance, Holm-Bonferroni corrected. * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, *ns* indicates $p > 0.05$.

Key points	L1	L2	L3	DeepLabCut
Left Ear	0.86 (0.96, 0.78)	0.92 (0.87, 0.98)	0.93 (0.91, 0.94)	0.89 (0.91, 0.88)
Right Ear	0.88 (0.97, 0.81)	0.89 (0.83, 0.95)	0.90 (0.84, 0.96)	0.89 (0.90, 0.89)
Left Elbow	0.95 (0.97, 0.93)	0.95 (0.96, 0.94)	0.94 (0.93, 0.95)	0.94 (0.94, 0.93)
Right Elbow	0.96 (0.94, 0.97)	0.92 (0.90, 0.94)	0.92 (0.91, 0.94)	0.93 (0.94, 0.91)
Left Eye	0.92 (0.98, 0.87)	0.88 (0.81, 0.98)	0.94 (0.98, 0.90)	0.93 (0.92, 0.93)
Right Eye	0.92 (0.99, 0.87)	0.92 (0.85, 0.99)	0.94 (0.96, 0.93)	0.92 (0.94, 0.90)
Mouth	0.93 (0.98, 0.88)	0.91 (0.84, 0.98)	0.95 (0.98, 0.91)	0.92 (0.92, 0.93)
Nose	0.95 (0.98, 0.92)	0.93 (0.88, 0.98)	0.96 (0.98, 0.94)	0.93 (0.93, 0.93)
Left Shoulder	0.80 (0.98, 0.68)	0.83 (0.72, 0.99)	0.94 (0.93, 0.95)	0.88 (0.90, 0.86)
Right Shoulder	0.85 (0.97, 0.76)	0.82 (0.72, 0.96)	0.91 (0.90, 0.93)	0.88 (0.90, 0.87)
Left Wrist	0.97 (0.98, 0.96)	0.95 (0.95, 0.94)	0.95 (0.97, 0.94)	0.94 (0.96, 0.93)
Right Wrist	0.95 (0.94, 0.96)	0.93 (0.90, 0.95)	0.93 (0.93, 0.93)	0.93 (0.94, 0.91)

Table 3.1 Precision, Recall, and F1 scores. Results of each labeler and DeepLabCut algorithm against consensus annotation, for each key point. Results are given in the form $\frac{\text{F1 score}}{(\text{Precision}, \text{Recall})}$.

Key points	L1	L2	L3	DeepLabCut
Left Ear	3.49	3.06	2.93	3.49
Right Ear	2.76	3.04	3.07	3.39
Left Elbow	2.56	2.91	2.60	3.05
Right Elbow	2.74	2.39	2.71	2.89
Left Eye	3.15	2.89	3.09	3.36
Right Eye	2.74	2.75	2.84	3.06
Mouth	2.68	2.58	2.42	3.08
Nose	2.18	2.58	2.30	3.05
Left Shoulder	2.59	3.63	3.23	3.22
Right Shoulder	2.64	2.80	2.71	2.87
Left Wrist	2.66	2.79	2.54	3.30
Right Wrist	2.67	2.67	2.86	3.23

Table 3.2 RMSE values. Averaged results of each labeler and DeepLabCut algorithm against consensus annotation, for each key point.

3.4 Comparison to a fourth human labeler

Aim of the task. To mitigate the bias from subjective feature annotations, we trained our network on a consensus annotation obtained by pooling the annotation of three individual experts. Next, we compared the performance of our network and that of each individual expert to the consensus annotation. However, this comparison is unfair to the network because the annotations of individual experts are used to build the consensus. Thus, there is a risk of overestimating the reliability of human performance. A fairer comparison would be to compare the performance of the algorithm with that of an individual labeler whose data were not used to build the consensus. To do so, we recruited a fourth coder and asked her to label a subset of 250 frames within the entire dataset of 2000 frames. Frames were randomly selected to reflect a similar percentage within each fetus and gestational period. The fourth labeler, instructed to identify body parts in the same principled and consistent way as the previous three labelers (see par. 3.2.1.3), annotated

12 fetal anatomical landmarks. The human annotation time of the subset of frames was estimated at around 4 hours.

Data analysis. The performance of the fourth coder in assigning labels was compared against the *consensus annotation* previously generated from the annotation of the three labelers. The performance was then evaluated using the two accuracy metrics described in par. 3.2.1.5 (i.e. the F1 score and the root mean square error, RMSE). The algorithm's performance on the frames' subset of interest was still available from the main study.

Results: comparing the fourth labeler against DeepLabCut.

F1 score. The F1 score of the fourth labeler was computed against the *consensus annotation*. Mean performances across body parts of the fourth labeler (L4) and DeepLabCut are shown in Table 3.3. L4 had an average performance of 0.86 ± 0.06 , while DLC showed a slightly higher average performance of 0.91 ± 0.02 . It should be noted that the L4 detection accuracy is also lower than the three labelers' accuracies obtained in the main study (in the range of 0.90 – 0.93, see par. 3.3.1).

RMSE. The RMSE of position estimation by L4 was computed against the *consensus annotation*, for the true positive (TP) detections. The mean RMSE across body parts by L4 was 4.17 ± 0.57 (Table 3.3), while DLC showed a value of 3.12 ± 0.28 . The mean delta RMSE for DLC was thus 1.05 pixels smaller than that of L4.

Performer	F1 score	RMSE
L4	0.86 ± 0.06	4.17 ± 0.57
DeepLabCut	0.91 ± 0.02	3.12 ± 0.28

Table 3.3 F1 scores and RMSE values of the 250 frames subset. Averaged results (mean $\pm$ standard deviation) of the fourth labeler (L4) and DeepLabCut algorithm against consensus annotation, across key points.

3.5 Use case: extraction of wrist velocity

Aim of the task. One main use of automated pose estimation algorithms is to perform kinematic analysis. An important validation step is thus to show that kinematic variables extracted based on DLC compare well kinematic variables extracted based on annotations of human labelers. With this in mind, in this analysis we focused on wrist velocity. The logic was to assess whether the reconstruction of the wrist velocity profile is comparable between the algorithm and the human coder. We focused on wrist velocity because inspection of the bell-shaped profile of wrist velocity has been used previously to characterize prospective planning of reaching movements to mouth, eyes (Zoia et al., 2013; Zoia et al., 2007) and co-twin (Castiello et al., 2010).

Movement identification. Eighteen relevant movements (i.e. 13 from fetuses in the 1st trimester; 5 from fetuses in the 2nd trimester fetuses) of the left fetus' wrist were isolated by an expert user. Fetal movements of interest were i) hand-to-face movements and ii) non-targeted movements. Movements were selected as follows: i) the hand is clearly visible throughout the sequence; ii) an isolated arm movement aiming either towards the face or away from the body is detectable. Identifying the starting and ending areas can be challenging although not crucial to the purpose of the use case.

Frames annotation. To create a set of annotated movements, we extracted all the frames across the time course of the movement sequences. A trained coder labeled frames following the criteria mentioned in par. 3.2.1.3 thus generating position coordinates of labels in each frame.

Video analysis. Movement sequences were analyzed with DeepLabCut and predicted label coordinates were generated.

Kinematic analysis. Variation in anthropometric parameters across fetuses and within the same fetus at different gestational weeks precludes the use of an absolute measurement unit (e.g., millimeters). To compare the amplitude of the movements across fetuses we adopted the ulna length (i.e. distance between

wrist and elbow markers) as a relative measure. This has been supported by evidence relating fetal ulna length and gestational age (Chitty & Altman, 2016).

Data analysis. The wrist velocity was defined as the module of the velocity of the wrist marker. Wrist velocity was then normalized by the ulna length. The procedure was applied to the wrist coordinates extracted based on DeepLabCut and the human coder, per movement sequence (i.e. 18 sequences). Mean velocity of the wrist was computed across movement sequences on DeepLabCut and human annotations. A paired sample t-test (two-tailed) was then used to compare average values of wrist velocity for each sequence obtained based on DeepLabCut to the average values of wrist velocity extracted by human annotations.

Results. Across movement sequences, the mean velocity of fetal wrist extracted by DeepLabCut was 1.32 ± 0.16 (mean $\pm$ SEM), and the mean velocity of wrist obtained by human annotations was 1.38 ± 0.17 (mean $\pm$ SEM). Paired sample t-test revealed that average values of wrist velocity computed based on DeepLabCut did not differ from the average values of wrist velocity extracted by the human ($p = 0.36$).

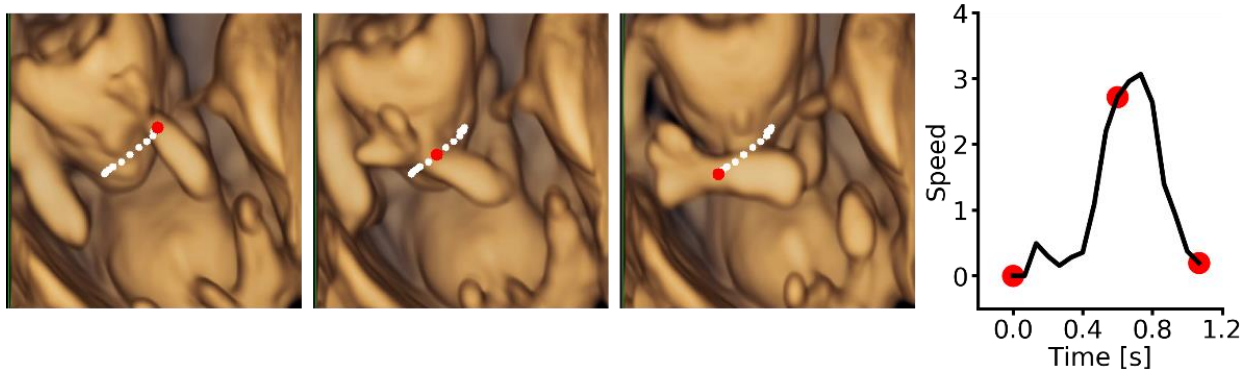


Figure 3.4 Example of trajectory and velocity profiles based on DeepLabCut predictions. The trajectory of the left wrist is computed for a representative movement performed away from the body in a 13-week fetus recorded at 15 hz. Trajectory of the wrist throughout the movement is depicted by white dots in three example frames in the early, middle and late phases of the movement. On the right is shown the velocity profile of the movement where the three timestamps shaping a bell-curved velocity profile were marked.

3.6 Discussion

Behavior is a dynamic phenomenon that involves changes to a fetus' pose over time. Reliable descriptions of human fetal behavior and movement may provide an opportunity to quantify the motor output of the fetal brain. However, most studies reported qualitative and (too) subjective descriptions of fetal behavior. In this chapter, we introduce markerless pose estimation of user-defined body parts, using deep learning architectures, as an effective tool for extracting key points of fetuses from video recordings. To quantitatively evaluate the performance of DeepLabCut, we benchmarked its results against manual annotations. Since it is not possible to establish an objectively defined ground truth in fetal 4D datasets, we collected manual annotations from 3 independent labelers. The three users' annotations were then combined to generate a final *consensus annotation*, which was taken as a reference dataset to benchmark the algorithm against human annotations on our US dataset. The combination of the various annotations led to a more reliable set of labels and also quantifies the limits of human performance. The deep network was validated on a dataset of 2000 frames containing labeled body parts of 37 different fetuses in 1st trimester of gestation, with a re-examination of 12 of them in 2nd trimester. Frames were extracted from videos resulting from high temporal resolution 4D US scans (i.e. average of 16 hz in 1st trimester fetuses and 11 hz in 2nd trimester). This was done to evaluate network performances even on relatively low-quality images, assuming that future studies will need high frame rate videos to perform reliable measurements of fetal movements.

Performances of DeepLabCut and labelers against *consensus annotation* were quantified by using two measures: a label detection accuracy measure (F1 score) and a position estimation error between human labels and consensual position of labels (RMSE). For both measures, the algorithm achieved close-to-human performance. By comparing DeepLabCut and labelers we found a mean delta value of 0.4 pixels (i.e. RMSE), reflecting the algorithm's high spatial accuracy.

A strength of our study is highlighted by the similar performances shown by the three users. Both F1 scores and RMSE values, with a few exceptions on some body parts, are consistent between the human

coders. In our opinion, this reflects the systematic and standardized training described in par 3.2.1.3 that coders received before starting the labeling procedure. Moreover, this point further strengthens the performance of the algorithm which competed against humans with strong expertise in performing the specific task.

Comparison to a fourth human labeler. Considering the unfair comparison for the network due to the involvement of the individual human annotations in the building of *consensus annotation* we evaluated whether and to what extent the reliability of human performance was overestimated. To test this, we carried out a further task by comparing DeepLabCut against a fourth human labeler whose data were not used to build the *consensus annotation*. The fourth trained coder labeled a random subset of 250 frames and the performance was then compared against the *consensus annotation* (for the same subset of frames) previously generated by the three labelers. Detection accuracies of the fourth coder were found to be slightly lower than F1 scores of DeepLabCut, as well as slightly lower than the three labelers involved in the *consensus annotation*. This is confirmed by the RMSE values, showing the same patterns between L4, DeepLabCut, and the three labelers. The result could be explained by the reduced learning of the fourth coder in assigning labels. This is not surprising when considering the proficiency at a task as increasing with increased experience. The lower performance of the fourth coder could thus reflect a shorter amount of experience (i.e. annotation of 250 frames), compared to the acquired experience of the three coders throughout the task (i.e. annotating 2000 frames).

Use case. A use case of our network is provided to validate its potential in kinematic analysis, a milestone in the study of fetal movement. The task aimed to compare the velocity of the fetus' wrist extracted based on DeepLabCut and humans, to assess consistency between performers in extracting kinematic variables. To do this, eighteen movements of interest were identified by an expert user, then frames were labeled along the entire sequences. The profile of the left wrist marker predicted by DeepLabCut and annotated by the coder was computed for the selected movements, then compared. The comparison revealed no statistical difference, confirming the robustness of using DeepLabCut for kinematic tasks. As remarked

earlier, manual annotations were performed by a human coder systematically trained on the labeling of the fetal body part in 4D US imaging. This makes the task for the algorithm even more challenging.

The broader purpose of this chapter was to investigate fetal movement and behavior using novel methods to make fine motion tracking of fetal body parts. The annotated dataset we generated and described through the help of a pose estimation algorithm, as well as wrist velocity computing, is just the beginning. The next step will be to perform a careful and more comprehensive analysis of fetal upper limb kinematics. Once extracted reliable movement parameters, multiple research questions should be investigated. For example, findings reported by studies assessing the kinematics of fetal reaching movements should be replicated by collecting US scans at higher temporal resolution and applying automated systems. Motion tracking tools should also be used to consistently segment movements. Starting from such standardized practices, the existence of a prospective control in human fetuses should be explored. Do fetuses already show anticipation (Delafield-Butt & Gangopadhyay, 2013)? Do fetuses prospectively calibrate their reaching patterning to the property of the target (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007)? The presence of an anticipatory structure in fetuses has already been suggested, however, this hypothesis needs to be confirmed within a quantitative and reliable framework. A second fascinating research question concerns the individual style of movement. Evidence showed that outputs of the motor system may vary substantially from one individual to another (Ting et al., 2015). Individual variations in movement patterns have been documented in a wide variety of animals (Calabrese et al., 2011), as well as in humans (Koul et al., 2016; Welch & Ting, 2008). Can we identify individual movement styles in human fetuses? Whether this is detected, is it consistent across different gestational periods (e.g., examining fetuses at around 13 weeks, then around 20 weeks of gestation)? Moreover, fetal movements in children who were later diagnosed with autism spectrum disorder (ASD) have not yet been studied. However, one could speculate that postnatal movement patterns of autistic children may be present in the womb (Gamliel 2012). By providing an effective tool to extract movement parameters, Chapter 3 stands as an essential first step in performing a careful analysis of fetal movement

kinematics. An automated and reliable assessment of fetal behavior has thus the potential to detect abnormal motor patterns of development in utero.

Here, we provide a benchmark of DeepLabCut on multiple manual annotations. The trained network, as well as the dataset with consensual annotations, will be released to the scientific community and this allows the user to bypass the manual labeling and the training process. Currently, DeepLabCut trained on this dataset can be used directly for inference on novel videos. Users will then be allowed to use the trained network to make pose estimation predictions on their own datasets, collected according to their specific task needs.

Analyzing video with DeepLabCut trained on this dataset would overtake two issues of classical studies reporting on the kinematics of fetal movements (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007). First, by eliminating the element of time required by manual coding. Then, by reducing the element of subjectivity and guaranteeing consistent segmentation across users. In the future, having the ability to skip labeling, training and run video inference with robust models will lead to more reproducible and scalable research.

3.7 Conclusion

We generated the first annotated dataset of fetal movements by using a 2D video-based markerless motion capture (i.e., DeepLabCut). The dataset was first manually annotated by 3 independent labelers that were instructed to locate body parts of fetuses in a consistent way. In a subsequent stage, annotations were combined to create a *consensus annotation* used to benchmark DeepLabCut and quantify the limits of human performance. The algorithm’s performance approaches the human performance, suggesting that our trained network succeeds in tracking fetal movement from high resolution 4D US imaging. The trained network will be released to the community and users will be allowed to apply it to their specific datasets, to generate and extract the desired movement parameters.

3.8 Author Contribution

Novella Pretti contributed to the conception and design of the study; supervised and performed the training of the human labelers; performed the extraction of frames to create the training dataset, frames labeled by the fourth coder and the movement sequences used for the wrist velocity computing; labeled the training dataset and the movement sequences used for the wrist velocity computing; contributed to design the analysis.

Chapter 4 - Fetal Visual Perception and Biological Motion Preference

4.1 Introduction

4.1.1 Prenatal visual development

Analysis of human brain development at an early fetal stage in utero offers unique insights into the prenatal sensory system. In the third trimester of pregnancy, perceptual information is processed in the fetal brain (Anderson & Thomason, 2013; Eswaran et al., 2004). Postmortem analysis of the human eye has revealed significant biological development from mid-gestation through to birth, with many of the essential components for visual function present starting around 25 weeks gestational age (GA) (Hendrickson et al., 2012). This retrospective observational study also described more advanced development in peripheral visual regions. Before post-natal development, peripheral vision appears to be more sensitive than foveal vision to detect environmental change. Research on visual function in preterm infants has shown that visual attention and fixation are present starting from 30 weeks (Morante et al., 2008; Romeo et al., 2012), although light stimulation can sometimes elicit physiological responses and evoke cortical potentials even in 28-week fetuses (Eswaran et al., 2004).

A modeling study has indicated a significant amount of luminance within the uterus, showing that fetuses in the third trimester develop in conditions that allow for visual experience before birth (Del Giudice, 2011). Moreover, animal studies have shown that light within the uterus is necessary for the typical development of the formation of the mammalian eye (Rao et al., 2013). Taken together, this evidence suggests that, starting from 28 weeks, fetuses are able to see in principle when exposed to sufficient lighting, even though with limited acuity. Furthermore, outlining a complex sensory environment for the fetus, these studies suggest that prenatal experience may be crucial for perceptual and cognitive development. However, little is known about the cognitive and perceptual abilities of fetuses.

In the visual domain, most studies have investigated the visual functional processing in human fetuses using simple visual stimuli such as flashes of light (Eswaran et al., 2004; Eswaran et al., 2002; Fulford et al., 2003; Matuz et al., 2012; McCubbin et al., 2007; Sheridan et al., 2008). Thus, very little is known about the ability of fetuses to process the percept of an image. This gap is mainly because of the difficulty of delivering visual stimuli inside the womb. Some recent studies have tried to fill this gap by projecting light stimuli in a face like versus non-face like configuration and comparing the behavioral responses of fetuses to these stimuli (Reid et al., 2017; Reissland et al., 2020). Using this approach, Reid and colleagues (Reid et al., 2017), showed in 34-week fetuses a preference to follow a top-heavy (i.e. face like) rather than inverted (i.e. non-face like) shape. A few years later, Reissland and colleagues (Reissland et al., 2020) extended this paradigm to three slightly more complex types of light stimuli (face-like top-and bottom-heavy designs, and a non-face-like control). The difference in relative head turns when exposed to face-like stimuli (both top-and bottom-heavy) compared to the control stimulus was significant, although head turns to top-heavy and bottom-heavy face-like configurations were similar.

The two studies observed fetal head turns to follow stimuli using four-dimensional ultrasound (4D US) technology. In the present study, we examine fetal responses using a custom-made visual stimulator (hereafter TCFVS, Transcutaneous Fetal Visual Stimulator) in combination with two-dimensional ultrasound (2D US). Our motivation for using 2D US rather than 4D US was that eye movements are difficult to detect from 4D US (Donovan et al., 2020). In 2D US, the lens of the eyes can be distinguished as white circles within the hypoechoic eyeball (see Figure 4.1) and both eye and head movements can be relatively easily detected (Inoue et al., 1986).

4.1.2 Why biological motion

Converging evidence suggests that preferential attention to biological motion (i.e., the movements of vertebrate animals) is a highly robust, evolutionarily well-conserved, and early-emerging mechanism of adaptive action. A spontaneous preference for looking at biological motion is observed within the first days of life in human neonates (Bardi et al., 2011; Bidet-Ildei et al., 2014; Craighero et al., 2017; de Klerk et al., 2016; Simion et al., 2008), and observed even upon first exposure in visually inexperienced newly-hatched chicks (Burr et al., 2005; Vallortigara & Regolin, 2006). Perception of and preferential looking towards biological motion can be elicited by a variety of forms (including the motion of non-human species as in (Simion et al., 2008)), is remarkably robust to signal degradation (Johansson, 1973), and is spared after early visual deprivation, even when global motion processing is impaired (Hadad et al., 2012). Together, these features suggest a biological mechanism that is evolutionarily critical for survival. Moreover, visual preference for biological motion can also be considered for its critical role in many aspects of social and communicative development (Pavlova, 2011). The criticality of attention to biological motion in early social development is underscored by the finding that this preference is disrupted in newborns with a high familial risk for Autism Spectrum Disorder (ASD) (Di Giorgio et al., 2016), and in 24-month-old toddlers diagnosed with ASD (Klin & Jones, 2008; Klin et al., 2009).

A seminal study by (Simion et al., 2008) provided the first demonstration that human neonates display a spontaneous visual preference for point-light displays of non-conspecific upright biological motion from the very first days of life. Several studies have since replicated and extended this finding, establishing that 2- to 4-day old human neonates exhibit preferences for the motion of both human and non-human species compared to inverted, random, and rigid non-biological motion displays (Bardi et al., 2011, 2014; Bidet-Ildei et al., 2014).

4.1.3 Our pilot study

A predisposition for biological motion has been reported in 2-day-old newborns (Bardi et al., 2011, 2014; Bidet-Ildei et al., 2014; Simion et al., 2008). Our pilot study capitalizes on this set of research to design a novel paradigm to deliver biological motion stimuli in uterus. The broader ambition, entrusted to a future study, is to use this paradigm to investigate whether a preference for biological movement already exists before birth. This chapter reports on the development and use of a custom-made visual stimulator (TCFVS) to deliver stimuli through the maternal abdomen. The TCFVS can display dynamic light configurations with biological and non-biological velocity profiles, adding new possibilities to the available paradigms using simple and static stimuli made of three LEDs (Reid et al., 2017; Reissland et al., 2020). In the current study, we used the stimulator in combination with 2D US to examine responses by third-trimester fetuses (28 to 32 weeks of gestation) to biological versus rigid motions.

Our primary aim was to validate this paradigm in a proof-of-principle study with 7 fetuses. The present pilot was designed to develop a new methodology to investigate fetal eye movements through the detection of the eye lens in 2D US scans. From this observation, we then derived two measures of fetal responses: the proportion of orienting responses and the looking times towards and away from biological and rigid motion stimuli. Looking time is commonly used in infant studies to measure how long infants engage with a stimulus. Differential looking times indicate discriminative processing and thereby provide clues to perceptual or cognitive abilities (Oakes, 2010). Here, we adapted the looking time paradigm for measuring the duration of head turns and eye gazes towards and away from visual stimuli.

Having assessed the effectiveness of the TCFVS in delivering stimuli and validated the looking time paradigm applied to fetal eye movements, we used the paradigm to obtain pilot data for data exploration and hypothesis generation. First, the hypothesis of a preference for biological motion predicts a larger proportion of orienting responses and a longer looking time towards visual stimulus displaying biological movement pattern compared to visual stimuli displaying a rigid movement pattern. Second, the proportion of orienting responses should be larger and the looking time longer towards biological stimuli

compared to away from biological stimuli. In the following, I present data from a pilot sample of seven fetuses.

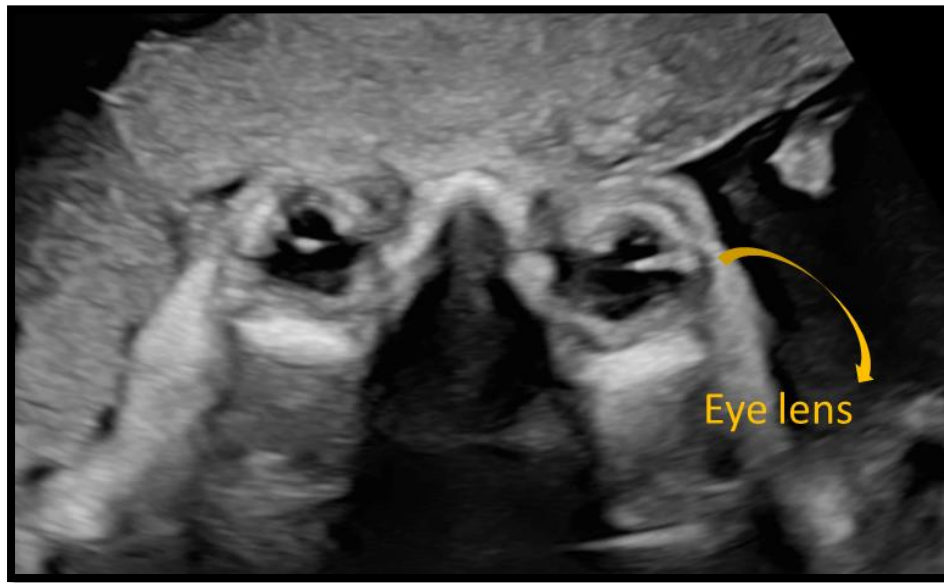


Figure 4.1 Coronal section of the fetal head at the orbital level in a 30-week-old fetus where the lens of the eyes (white circles) are easily visualized.

4.2 Methods

4.2.1 Participants

The research protocol was approved by the local ethics committee (ASL3 Genovese) and complied with the principles of the revised Helsinki Declaration (World Medical, 2013). All pregnant women participating received written information and gave their informed written consent to participate in the study.

All women had pregnancies with no known complications and a healthy BMI between 18.5 and 24.9. Behavioral responses to stimuli were recorded in fetuses of gestational age between 28 and 32 weeks ($M = 30$ w., $SD \pm 1.3$). An initial 2D scan informs about the fetal behavioral state and determines the eligibility of the fetus for recruitment. Fetuses in a “quiet awake” state (i.e. state 3F), assessed through the presence of eye movements and the absence of gross body movements (Nijhuis et al., 1982), were included in the final sample.

4.2.2 Data Acquisition

Ultrasound scans were performed at Gaslini Hospital using a GE Healthcare Voluson Ultrasound System E10 with 2D convex transducer C1-5. The scans were recorded and saved to an external hard drive for offline analysis. All scanning was performed by an experienced sonographer. Initially, routine fetal biometry measurements were collected to demonstrate normal growth and no anomalies.

4.2.3 Apparatus and Stimuli

4.2.3.1 Light source

To test fetal responses to light, we developed a custom-made noninvasive visual stimulator (i.e., the TCFVS) generating patterned light stimuli through a movable LEDs matrix. TCFVS interfaced with a customized laptop and is controlled with an AC medical grade power supply unit. The display is assembled from a Light Emitting Diodes (LEDs) matrix placed inside a translucent plastic enclosure (9 x 10 cm), Figure 4.2.A. Shape, dimensions, and functionalities of the TCFVS display have been specifically conceptualized and designed for applications with pregnant women in the hospital setting. In particular, the external case, in contact with the mother’s abdomen, is designed with a parabola shape to ensure maximum light dispersion through the skin. All components in contact with the skin are conceived in a translucent biocompatible material.

LEDs (Kingbright, Farnell) are stacked on a flexible PCB, with a resolution of 16x20 pixels (320 total pixels) and a surface area of 50.4 cm². Each LED is sized 1.2 mm, emitting at 640 nm with power of 0.8mW. LEDs are red-colored since the red portion of the spectrum allows the immature visual system a better discriminative performance (Adams et al., 1994; Del Giudice, 2011). The pitch (i.e. distance between pixels) is 2 mm on the horizontal axis and 3 mm on the vertical axis. Such a short pitch between LEDs is essential for an adequate reconstruction of the biological movement profile. Stimuli designed for the current experiment were shaped using a subset of 24 LEDs arranged in 6 dots, giving an external light illuminance on the participant's abdomen around 4000 lx.

Stimuli projected by the TCFVS result from data captured by an optical motion capture system (Vicon Motion Systems Ltd., UK) equipped with eight infrared cameras. To do so, a human participant wore 6 retroreflective-marker (6.5 mm in diameter) on the right hand and perform a hand movement (see par. 4.2.3.2.). Motion capture coordinates are converted to .avi files (i.e. 6 white dots moving on a black background, as shown in Figure 4.2.B) through the Real-Time Motion Capture Toolbox (RTMocap) in MATLAB (MathWorks) environment. Then, an in-house software toolchain converts the .avi files to visual stimuli displayed in the TCFVS.

The TCFVS device is complemented by a User Interface (UI) supporting the stimuli selection for the experiment, as well as the start and the stop of the device and stimuli. Moreover, the UI allows users to monitor the device status and to create an external timestamp logfile returning the time at which each event was generated, confirming the correct stimulation sequence along the experimental trials.

Before starting recruiting, the TCFVS was tested to verify the conformity with both the safety prescriptions and the electromagnetic compatibility requirements of reference documents.

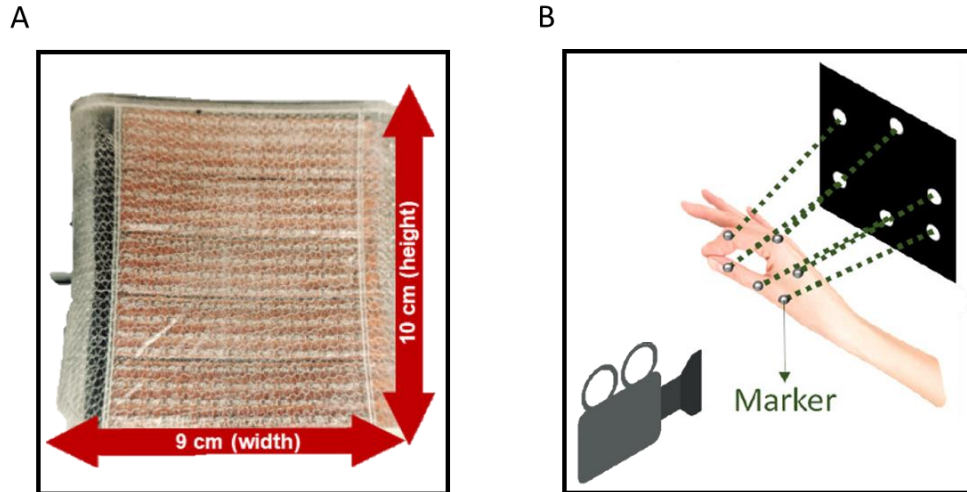


Figure 4.2 A) LEDs matrix display generating visual stimuli. B) Biological movement profile recording with an optical motion capture system.

4.2.3.2 Light stimulus configuration

Fetal behavioral responses were assessed by projecting on the maternal abdomen two types of light stimuli: a biological motion pattern (depicting the opening and closing of human hand's index and thumb) and a non-biological, rigid motion pattern. The biological motion stimulus results from coordinates acquired with a motion capture system, as described in par. 4.2.3.1. Ten hand opening and closing cycles were displayed in each trial (20s lasting) with a frame rate of 17.5 hz (Figure 4.3.A). The choice of this stimulus was driven by our focus on the velocity profile. The selected movement (i.e. hand opening and closing) contains an inherent characteristic of the human motor repertoire that is smoothness. According to the minimum hand jerk model, the biological motion has a characteristic velocity profile which is a cost function that minimizes jerkiness over a specified movement trajectory (Flash & Hogan, 1985). Moreover, constraints imposed by the display size limited the range of a potential target. The adaptation to the LEDs matrix of a modified form of point-light human hand movement seemed to be the best chance to provide a stimulus with properties of biological motion not shared by the rigid motion of an identical array of dots.

The rigid motion stimulus consisted of the rotation of one static frame from the hand animation (i.e. the maximum aperture) around its vertical axis, to produce the motion of a rotating rigid hand-like object. The rigid motion design is consistent with newborn's research, showing 2-day-old babies' preference for biological motion point-light display when contrasted with the rigid motion display (Bardi et al., 2011). Each trial (20s lasting) displayed the stimulus at a rate of 1 frame/s (Figure 4.3.B). The two selected stimuli (i.e. biological and rigid patterns) differed in terms of their velocity profile.

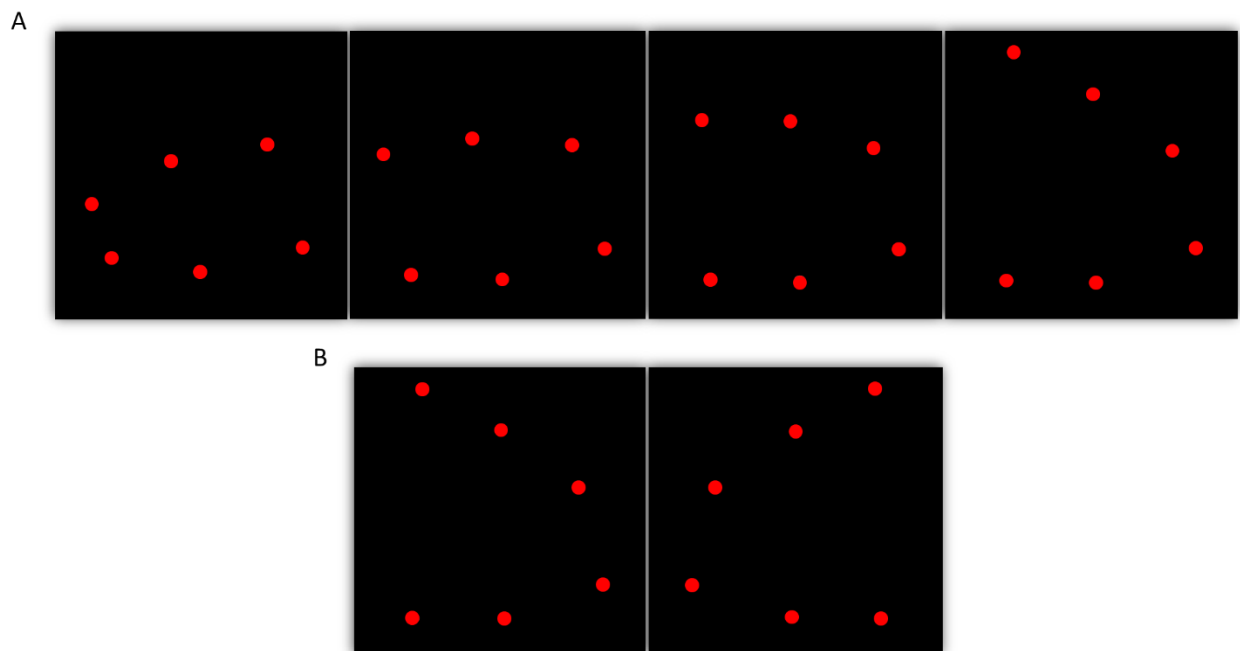


Figure 4.3 Schematic representation of experimental stimuli. A) Selected frames sequence depicting the biological motion pattern. B) Frames that, alternating, shape the rigid motion stimulus.

4.2.4 Experimental design

All fetuses underwent 14 trials, each lasting 20 seconds, in which the fetal response was recorded. Ten trials were conducted by delivering a visual stimulation (i.e. 5 trials with random motion stimulus, then 5 trials with rigid motion stimulus), whereas in 4 trials the stimulation was absent (i.e. baseline trials). Moreover, to maximize the possibility of getting fetuses' attention, two stimulation modes were applied: i) a constant light presentation, where the experimenter holds the display steady next to the probe; ii) an intermittent light presentation, where the display was moved across the maternal abdomen in a horizontal direction away from the probe. We thus obtain a total recording time of 200 seconds (3.3 minutes) while stimulating and 80 seconds (1.3 minutes) of baseline periods. The sequence of trials was organized according to the order shown in Table 4.1. Both the initial stimulus type (biological-rigid motion) and the stimulation mode (intermittent-constant) order of presentation were counterbalanced across subjects.

Table 4.1 Experimental trial sequence

Trial n.	Stimulus Type	Light presentation mode
1	No stimulation	NA
2	No stimulation	NA
3	Biological motion	Intermittent
4	Biological motion	Intermittent
5	Biological motion	Intermittent
6	Biological motion	Constant
7	Biological motion	Constant
8	No stimulation	NA
9	No stimulation	NA
10	Rigid motion	Intermittent
11	Rigid motion	Intermittent
12	Rigid motion	Intermittent
13	Rigid motion	Constant
14	Rigid motion	Constant

NA: not applicable

4.2.5 Procedure

An initial set of fetal biometry measurements through 2D US were taken by the ultrasonographer with the twofold purpose of demonstrating normal fetal growth and assessing the fetal behavioral state. Pregnant women, lying on their back, were then asked not to talk during the study and to remain as still as possible to optimize image quality. The 2D scan also informed the experimenter of the precise location of the fetal head before the presentation of the stimuli. This gave the greatest likelihood that the experimenter could place the light in the peripheral visual field of the fetus. Stimuli were delivered on the maternal abdomen to the side of the fetal face (left, $n = 2$; right, $n = 5$), relative to the position of the fetus (all fetuses were in cephalic position/head down). The experimenter, not blind to the condition, took advantage of the ultrasound image for initial placement but then attended to the movement of the stimulus across the maternal abdomen, neglecting the viewing of the ultrasound image once the trial has started. Stimulus delivery was standardized across conditions ensuring equivalent timing between conditions. The TCFVS device is controlled by the experimenter via a laptop. After sending the command for the selected stimulus the experimenter, depending on the trial, holds the display steady next to the probe or moved the display across the maternal abdomen in a horizontal direction away from the fetal central visual location, for 5 s. This speed takes into account constraints specific to this population, such as the width of maternal tissue accessible to present stimuli and the space available for the fetus to move (Reid et al., 2017). US recording, light stimuli, and movement of the display across the abdomen (whether required by the trial) start simultaneously after a voice signal ‘go’ and automatically stop after 20 seconds. This procedure was repeated 10 times (i.e. 5 times for each stimulus). Moreover, 4 total baseline scans were recorded. Participants were blind to conditions. The stimuli presentation was consistent with that previously reported in (Reid et al., 2017).

4.2.6 Data coding

During scan recording, the ultrasound image was rotated by the sonographer so that the fetal face appeared upright. This was done to simplify the coding at a later stage, consisting of simply *left* and *right* head/eye movements in the coronal plane, taking account of fetal position in the womb. Coding of recordings was conducted offline using a frame by frame motion analysis software (i.e. Kinovea[®] software).

For each fetus, both the direction of the head/eye movements and the duration of movements away from the individual fetus' initial resting head/eye position were coded. At the start of each trial, the resting position of the eye lens on the ultrasound image was coded. At the onset of a head/eye movement (defined as the first frame where the movement occurs), the coder annotated the start time (i.e. frame number) and the direction of movement (i.e. left, right). Then, the offset of the movement (defined as the last frame where the movement occurs, after the head/eye return to its initial resting position) was also coded. The first coder, blind to the position and the type of stimulus, repeated this procedure for each movement that occurred within the same trial. The onset of any movement was not constrained to the onset of the stimulus but no movements were coded following stimulus offset. Whether the offset of a movement was not detectable, due to the end of a trial while the movement is occurring or to intermittent occlusion of the eye lens, a different code was assigned so that only the direction of this movement was included in the final analysis. Conversely, trials where the fetus performs gross up and down head movements, making the detection of eye lens impractical along the trial, were excluded both from the direction and the duration analysis of movement.

Following blind coding, *left* and *right* movements were transformed to indicate whether movements were made towards or *away* from the stimuli. An orientation towards the stimulus was a head/eye movement in accordance with the position of the light delivered during the trial. An orientation *away* from the stimulus was a head/eye movement performed in the opposite direction to the position of the light source.

The direction of the head and eye movements were subsequently analyzed frame by frame by a second blind coder (on 100% of the sample). Cohen's kappa was performed to determine interrater agreement on both the presence and direction of the first fetal movement within each scan, at a single-trial level. A substantial agreement between two coders' judgments was found, $k = 0.892$, $p = <0.001$. The behavioral measurement was therefore deemed appropriate for use in hypothesis testing in the current pilot study (Landis & Koch, 1977).

4.2.7 Data analysis

Fetal responses to light stimulation were defined as any head turn or eye movement observed during the presentation of the light stimuli. For each codable trial, two dependent variables were computed and analyzed: the proportion of orienting responses and the looking time.

The proportion of orienting responses was defined as the proportion of fetal responses towards the stimulus. Analyses of fetal behavioral responses were carried out based on the coding performed by the first coder. Looking time was defined as the amount of time the infant spent looking at the stimulus. To compute looking times, we first calculated the number of frames in which the eye or the head remains oriented in a given direction (*n_frames_movement*). Then, taking into account the frame rate of the video (FR), we computed looking time (in seconds) as the number of frames in which the eye or the head remains oriented in a given direction (*n_frames_movement*) divided by the frame rate of the video (FR):

$$looking\ time = (n_frames_movement)/FR$$

Values expressed in s were then converted in ms. Although all movements were coded across trials, only the first head or eye movement within each trial was included in analyses, in line with infant research (Gliga et al., 2009). This applies to both the orienting responses and the looking time computations.

Validation of methodology. Two analyses were performed to test the effectiveness of our visual stimulator (TCFVS) in delivering light stimuli that can actually be perceived by the fetus. First, to evaluate whether light stimulation had an effect on eliciting fetal movements, we compared the proportion of orienting responses and looking times during stimulation and baseline periods (i.e., no stimulation), regardless of the type of stimulus (biological versus rigid). Second, to test the hypothesis that light stimulation attracts fetuses' attention, we compared the proportion of orienting responses and looking times towards the stimuli and away from the stimuli, regardless of the stimulus type.

Preference for biological motion. To examine whether fetuses showed a preference for biological motion, we compared the proportion of orienting responses and looking times across conditions (i.e. towards_biological; towards_rigid; away_biological; away_rigid).

All fetuses presented normally distributed data. We reported mean and standard error of the mean (SEM) for each above-described condition (values are depicted in figures), whereas for all comparisons between dependent variables we estimated the effect size (d_z) with alpha set at 0.05, and power set at 0.95.

4.3 Results

Primary analyses to validate the paradigm revealed a greater proportion of orienting responses during stimulation compared to baseline trials (with an effect size of $d_z = 0.70$) (Figure 4.4.A). Moreover, the looking time was longer during stimulation trials compared to baseline trials, with an effect size of 0.56 (Figure 4.4.B.). This indicates that regardless of stimulus configuration, fetuses responded to visual stimulation. The proportion of orienting responses towards the stimulus was greater compared to the proportion of orienting responses away from the stimulus ($d_z = 0.60$) (Figure 4.5.A). Higher looking times were also

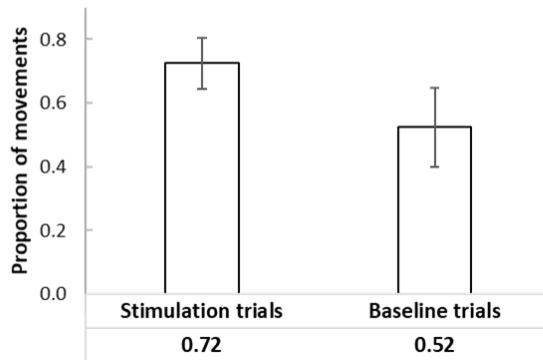
found towards the stimulus compared to away from the stimulus ($dz = 0.70$) (Figure 4.5.B). This indicates that regardless of the stimulus type, fetuses oriented their attention preferentially towards the stimulus.

Having verified the potential of visual stimuli to capture and maintain the attention of fetuses, we next compared the proportion of orienting movements and looking times across conditions (Figure 4.6.A.). We found that the proportion of orienting responses towards biological motion stimuli was greater than the proportion of orienting responses towards rigid motion stimuli ($dz = 0.33$). Moreover, the proportion of orienting responses towards biological stimuli was greater than the proportion of orienting responses away from biological stimuli ($dz = 0.82$). Looking times were also longer towards biological stimuli than towards rigid stimuli ($dz = 0.58$), and towards biological stimuli than away biological stimuli ($dz = 0.82$), Figure 4.6.B. If confirmed in the main study, these results would support the hypothesis of a fetal preference for biological motion.

Additionally, we found that the proportion of orienting responses away from the rigid motion stimuli was greater than the proportion of orienting responses away from the biological stimuli ($dz = 0.34$). Moreover, the proportions of orienting responses away from the rigid stimuli and towards the rigid stimuli were similar ($dz = 0.03$). A similar pattern was also observed for looking times.

Orienting responses were longer in the condition away from the rigid motion than away from the biological stimuli ($dz = 0.64$), and similar in the conditions away from the rigid stimuli and towards the rigid stimuli ($dz = 0.14$).

A



B

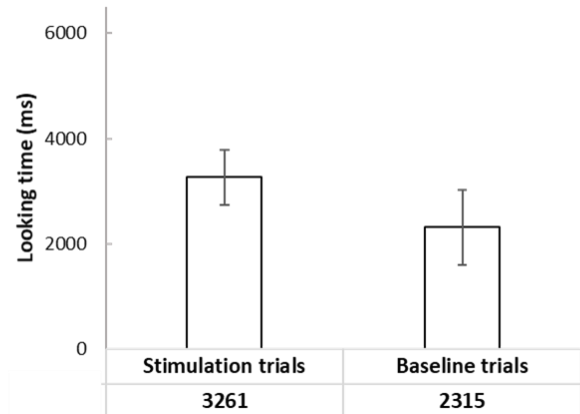
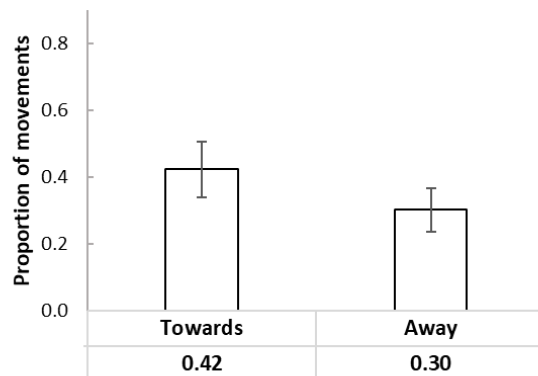


Figure 4.4 A) Mean proportion of movements performed during stimulation and baseline trials (i.e. no stimulation) across fetuses. Error bars indicate SEM. The table shows the mean values. B) Looking times (ms) during stimulation and baseline trials across fetuses. Error bars indicate SEM. The table shows the mean values.

A



B

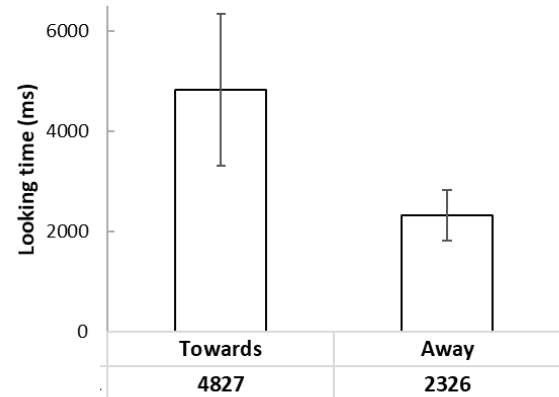


Figure 4.5 A) Mean proportion of fetal movements performed towards the stimuli and away from the stimuli, across fetuses. Error bars indicate SEM. The table shows the mean values. B) Looking times (ms) performed towards the stimuli and away from the stimuli, across fetuses. Error bars indicate SEM. The table shows the mean values.

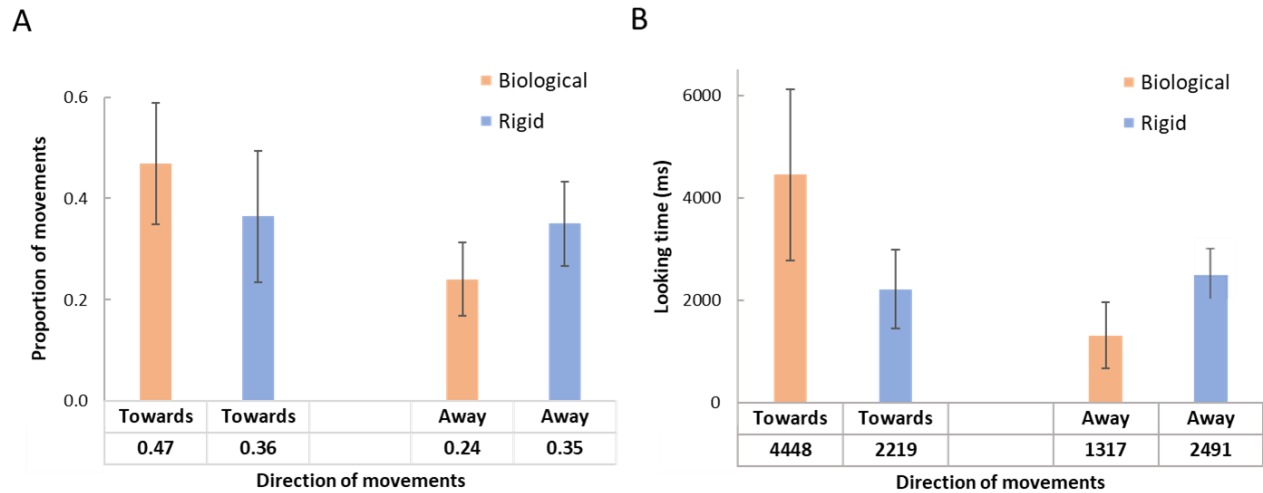


Figure 4.6 A) Proportion of orienting movements across fetuses compared per condition. Error bars indicate SEM. The table shows the mean values. B) Looking times (ms) across fetuses compared per condition. Error bars indicate SEM. The table shows the mean values.

4.4 Discussion

Existing evidence suggests that starting from 28 weeks, human fetuses are in principle able to see and process visual information (Del Giudice, 2011; Donovan et al., 2020; Eswaran et al., 2004). Furthermore, when exposed to light stimulation, fetuses are likely to modulate their behavioral response with respect to the stimulus configuration (Reid et al., 2017; Reissland et al., 2020). Well replicated findings reveal then that 2-day-old newborns show a spontaneous visual preference for biological movement profile. However, no study to date has investigated whether a similar preference exists in fetuses. This gap has arisen mainly because of the difficulty to deliver biological motion stimuli in uterus. Here we present a pilot study aimed at assessing the feasibility of a novel visual stimulation methodology designed for fetal research. Our first goal was to examine whether fetuses at 28-to-32 gestational weeks respond to light configurations delivered with a custom-made visual stimulator (TCFVS), as well as develop a looking time procedure based on fetal eye lens detection.

In our pilot sample of 7 fetuses, we found more fetal eye and head movements in response to light stimulation than in baseline periods, confirming that the optimized stimulation protocol produced an effect in eliciting behavioral responses. This is further supported by the higher fetal looking times found during stimulation trials. Our result is in line with the significantly higher number of fetal eye movements *during* light presentation than *before* reported in previous research (Donovan et al., 2020).

It could be argued that our light stimulation only induces a behavioral activation of the fetus due to the perturbation of the uterine environment. However, the greater proportion of movements performed towards the stimuli than away from the stimuli reveals a preferential behavior in the direction of the stimulus, suggesting that this possibility should have not occurred. Again, this is supported by the higher looking times found towards the stimuli, regardless of the configuration of the stimulus.

The analyses performed to pilot the methodology revealed large effect sizes. The behavior of our pilot fetuses seems to be strongly activated by light stimulations delivered with the TCFVS ($d_z = 0.70$) and straight to the lights ($d_z = 0.60$). Medium-to-large effect sizes were also found when repeating the analyses by computing the looking times, with longer looking times when stimulation was delivered ($d_z = 0.56$) and longer looking times when looking towards the stimulus. These results suggest that our stimulation protocol is effective in eliciting a response from fetuses and can be used to test specific predictions deriving from the hypothesis of a fetal preference for biological movement.

Looking at gaze behavior across stimulation conditions, revealed that fetuses in our sample responded more frequently towards the biological motion than towards the rigid motion (however, the small effect size should be noted), as well as more frequently towards the biological stimulus than in the opposite direction (with a large effect size). Similar patterns were found when measuring looking times across fetuses: looking times were higher towards the biological stimulus.

Surprisingly, fetal responses made away from the two stimuli revealed an inverse pattern. We found more responses away from the rigid motion than away from the biological motion, in line with fetal looking times. The pattern seems to reflect different processing of the rigid motion. One could speculate that when subtracting some inherent properties of biological motion as its smoothness, the prenatal visual

processing of movement could be less sensitive or negatively affected. Due to the evolutionary critical role of biological movement for survival, the fetus' sensitivity to its characteristics may have developed earlier.

Further, similar responses were made in reaction to rigid motion (both towards and away from the stimulus), this applied both to orienting responses and looking times. These findings seem to reflect a lack of modulation in fetal movements when delivering a rigid pattern. Conversely, a differential response (made toward or away) comes out strongly when presenting the biological motion. Whether and to what extent a biological velocity profile will trigger differential behavioral responses even before birth will be a crucial point to explore in the following full study.

Although the results of the present pilot study are in line with the preference for biological motion at birth shown elsewhere, paradigm and stimuli property are necessarily different from those adopted in infant research. When investigating the development of the human visual system with respect to the transition from fetus to infant, an increasing effort to optimize light stimuli and refine experimental procedures is essential for future studies. Those paradigms should also take into account the multiple difficulties arising when exploring visual processing in fetuses (Reid & Dunn, 2021). For example, the attrition rate of fetuses recruited in most studies is very high. Dunn and colleagues reported that approximately only 50% of the sample provides data that can be used in the analysis (Dunn et al., 2015). During the recruitment of fetuses involved in this pilot study, the rate seemed to be even lower. Approximately 1 out of 5 pregnant women giving the consensus to participate were then actually tested. The main barrier was the position of the fetus in the womb and the inability of modifying this initial state. The fetus was often positioned such that the light stimuli could not be presented. This happened when the face of the fetus was directed towards the mother's back or side, when the placenta occluded the fetal face, or when the face was occluded by the hand/s of the fetus. The eligibility of the fetus for testing was then determined by its behavioral state (Nijhuis et al., 1982). Fetuses in the “quiet sleep” state (i.e. state 1F), assessed through the absence of eye movements and the absence of body movements were not involved as judged to be asleep.

A broader aim of the study is to foster knowledge on the development of fetal sensory capacities and information processing. Here we established that it is possible to test specific fetal fine-grained reactions (i.e. eye movements) to light stimulation. A potential implication of this work figures the possibility that the lack of fetal reaction to specific lights could be used in the future as a prenatal diagnostic marker for Autism Spectrum Disorder (ASD) risk. The speculation afford on findings reporting an abnormal biological motion processing in children with ASD is disrupted in newborns with a high familial risk for Autism Spectrum Disorder (Di Giorgio et al., 2016) and infants diagnosed with ASD (Klin & Jones, 2008; Klin et al., 2009).

As a methodological novelty, in the present study we propose for the first time a looking-time paradigm to explore visual processing and attention mechanisms in fetuses. The average looking time elicited by stimuli might offer substantial insight into fetal development in utero that would otherwise not be possible to observe. This raises eye movements as a fine-grained measure of fetal behavior, which is likely to provide a direct measure of fetal attention to visual information (Reid & Dunn, 2021). Furthermore, to the best of our knowledge, there are no standard devices designed for fetal research capable of generating any type of light configuration, including the transposition of human movements captured by motion capture systems. Indeed, other studies designed rather simple stimuli, with only 3 (Reid et al., 2017) and 12 LEDs (Reissland et al., 2020) displaying static configurations. More research is needed to develop this pilot study and a full study serves to replicate our initial findings on the fetal preference for biological movement.

4.5 Conclusion

The first conclusion of this pilot study is that visual processing of stimuli occurs in utero, indeed, fetuses modulate their behavioral responses according to the presence (or not) of a light source. The effectiveness of our stimulation protocol is demonstrated, suggesting that a full study can be carried out with such a visual paradigm. Preliminary results testing the fetal preference for biological movement in a pilot sample

reveal promising directions to be investigated. Future studies will explore whether there is an innate sensitivity to some properties of biological motion which are not shared by the rigid motion of an identical array of dots.

Chapter 5 - Discussion

5.1 Summary of the main findings

The primary aim of this thesis was to develop a quantitative experimental and analytical framework to investigate fetal behavior development. Single variables of fetal behavior were here investigated. Specifically, we aimed at examining fetal behavioral variables by i) tracking fetal movement assessed with 4D US, and ii) evaluating fetal responsivity to visual stimulation through 2D US. This was done using different techniques including ultrasound, deep neural network, motion capture system, and attentional measures.

Below, I briefly summarize the aim and main findings of each chapter before discussing the clinical implications and outlooks of these results.

When reviewing literature on fetal movements, one is confronted with several challenges and limitations in quantifying and interpreting 4D US fetal movement data. The aim of Chapter 2 was to provide an overview of some of these challenges and, based on the results of a systematic review, to examine their impact on fetal movement research. Aggregating studies based on outcome measure (i.e., frequency of fetal movements, quality of fetal movements using KANET test, and fetal kinematics) revealed gaps and limitations. Two prominent concerns arise from the low temporal resolution of 4D US, and the subjectivity that examiners introduced in their observations. The Chapter concluded that addressing these gaps will require advances in both the technique (by increasing the frame rate of US imaging) and the methods, with an improvement of procedures, analyses, and standardization of reporting practices. Finally, the Chapter pointed to promising future directions in characterization of fetal behavior, highlighting the application of deep neural networks trained on high-frame rates.

Chapter 3 aimed to move beyond some of the limitations highlighted in Chapter 2, by developing a quantitative approach to make fine motion tracking of fetuses in the first and second trimesters. This work aimed to generate the first annotated dataset of fetal movements by expanding a markerless pose estimation algorithm (i.e., DeepLabCut) for high temporal resolution 4D US imaging. Developing an automated system for movement detection and tracking, indeed, would reduce human labor and guarantee consistent interpretations across users. To create the dataset, 12 fetal anatomical landmarks were identified to be manually annotated. Due to the low quality of rendered images, the operation of identifying fetal body parts in 4D US images can be challenging for human labelers. To address human variability in manual annotations, three independent labelers annotated the same dataset (i.e., 2000 frames). Then, the annotations were combined to create a consensus annotation and this was used to quantify the limits of human performance and to evaluate the performance of the algorithm. Detection accuracy (i.e., F1 score) and the error of position estimation (i.e., RMSE) between prediction of algorithm and consensual annotation were computed. By comparing DeepLabCut and labelers we found a mean delta value of 0.4 pixels (i.e. RMSE), reflecting the algorithm's high spatial accuracy. To additionally benchmark our network, we then compared its performance against that of the fourth human annotator whose annotations were not used to build the consensus annotation. Finally, we demonstrated the potential of our network for kinematic analysis of fetal movement by using it to compute wrist velocity. The network we optimized succeeds in tracking fetal movements from high-resolution 4D US images and will be released to the community to allow users to extract movement parameters according to their task requirements.

Finally, in Chapter 4, I explored visual processing in third-trimester fetuses. The broader ambition is to investigate whether a preference for biological movement already exists before birth, in line with well-replicated evidence showing a biological motion preference in 2-day-old newborns. To this aim, we first pilot a novel visual stimulation methodology designed for fetal research. Our primary goal was to validate this paradigm in a proof-of-principle study with 7 fetuses. Specifically by i) examining whether fetuses at 28-to-32 gestational weeks respond to light configurations delivered with a custom-made visual

stimulator (TCFVS); ii) developing a looking time procedure based on fetal eye lens detection. Then, we used the paradigm to obtain pilot data for data exploration and hypothesis generation. Stimulations were provided through a custom-made LEDs device able to generate any type of light configuration, including those with a biological velocity profile. Fetal head and eye movements made in response to stimuli were observed in fetuses by simultaneously recording behavior through 2D US. Behavioral responses of our pilot sample are strongly activated by light stimulations delivered with the TCFVS and straight to the lights, with looking times confirming these findings. Results demonstrate the effectiveness of our stimulation protocol and can be used to test specific predictions deriving from the hypothesis of a fetal preference for biological movement. Moreover, we found a greater proportion of orienting movements towards the stimulus with biological motion than towards the one with rigid motion. A similar result was found in looking times, with higher duration of movement towards the biological stimulus.

Collectively, the studies outlined in this thesis demonstrate the feasibility and potential of a quantitative approach to fetal behavior and movement. 4D US affords unprecedented opportunities for observing fetal movement dynamics. However, the existing literature reviewed in Chapter 2, both evaluating spontaneous movements and movements in reaction to external stimuli, reveals gaps and limitations that need to be addressed. For example, major progress will require technical advances to improve the temporal resolution of 4D US, in particular for what concerns the kinematic measurements of fetal movements. A promising direction to enhance the quantitative analysis of 4D US movement data (even at high frame rates) is the adoption of deep neural networks as a tool for video-based motion tracking. Generating objectively annotated datasets to provide reliable movement parameters is a required condition to later extract accurate and precise measures. The application of automated systems to track fetal movements is validated in Chapter 3. The high performance of our network makes it suitable to be deemed a tool of choice for upcoming fetal movement studies. No less important for improving the reliability and validity of research findings is the adoption of rigorous procedures for data collection and analysis, including operational definitions and standardized coding procedures. These largely unaddressed

methodological and technical issues raise concerns about the utility and credibility of 4D US studies of fetal behavior. Among the studies reviewed in Chapter 2, there are a few examining fetal responses to visual stimulations. However, difficulties in delivering stimulations in utero limit research on fetal visual perception. Chapter 4 demonstrated the potential of a new visual stimulation methodology using a custom-made LEDs visual stimulator to project light stimuli through the maternal abdomen. Moreover, by adapting a looking time paradigm from infancy research, fetal attention towards different types of stimuli was quantified.

By usefully inheriting some lessons from other domains, such as computational studies and postnatal developmental research, this thesis aims to push research towards more reliable and interpretable evidence-based procedures in characterization of fetal behavior.

5.2 Clinical considerations

The scientific interest in the fetus' brain and behavior has a long history (Kisilevsky & Low, 1998). Since the advent of ultrasound imaging, much has been learned about spontaneous fetal movement patterns and behavioral states. This knowledge has proven to be useful for understanding physiological phenomena that are of clinical importance for fetal surveillance, such as fetal heart rate (FHR) patterns and Doppler blood flow velocity profiles (Mulder & Visser, 2016). The diagnostic value of behavioral measures was underlined in clinical examinations with 2D US and several aberrations of the normal developmental course have been reported under varying clinical conditions (Mulder & Visser, 2016). These abnormalities involve delayed emergence of movements, changes in the quantity (extremely low or high occurrence rate) or the quality of motor performance, and abnormal development of fetal behavioral states. Knowledge of normal fetal behavior also has helped to identify and characterize alterations in behavioral performance and development during pregnancies complicated by disease, drug exposure, or a malformed fetus. However, prenatal prediction of neurologic outcomes based on altered fetal behavior remains difficult.

There is no universally accepted procedure to test CNS functioning in the fetus to date, although there are a number of tests in use (e.g., Biophysical profile score; Fetal Neurobehavior Coding System, etc). A unifying theory of fetal brain and behavior is lacking (Mulder & Visser, 2016). This is because we are still far from understanding the basic neural mechanisms and functional significance of fetal movements and states. So, do fetal brain development and behavioral development run parallel during gestation? Such important questions remain.

From this gap in clinical and scientific knowledge, studies gathered in the dissertation aimed at pushing forward fetal research by providing a quantitative framework to analyze components of fetal behavior (i.e. body and eye movements). Identification of measurable indicators of fetal functioning that reflect the development of the peripheral and central nervous system (CNS) is the primary goal of fetal neurobehavioral research (Vindla & James, 1995). As an indirect measure of central nervous system integrity (Prechtl & Einspieler, 1997), fetal spontaneous movements might represent a further key factor to assess fetal physiology and its well-being (Anderson & Thomason, 2013). In this view, tools for detecting and tracking movement thus become valuable to quantify the motor output of the brain. The study in Chapter 3 proposes an easy-to-use and well-validated technology for a more objective and reliable characterization of fetal behavior. Moreover, being automated, the tool makes the procedures more suitable for incorporation into routine clinical practice. We anticipate that by opening up new possibilities for quantifying fetal movement, this tool will give a boost in the study of fetal movements and hopefully help establish an objective fetal movement science.

Fetus' responsivity to stimulation might also play a role in understanding neurobehavioral functioning. The study in Chapter 4 aims to foster knowledge on the development of fetal sensory capacities and information processing. Our pilot study established that it is possible to measure behavioral responses to the perceptual content of visual stimuli, with important implications for the advancement of fetal psychology.

Future progress requires the involvement of multiple disciplines to gain a comprehensive understanding of the role played by fetal behavior in development (Lecanuet et al., 2013). In the future,

several questions should be raised: 1) How is prenatal behavioral development involved in modeling the central nervous system? 2) How can clinicians and researchers establish behavioral baselines to assess the well-being of normally developing fetuses? 3) How can psychologists and neonatologists best apply knowledge of fetal behavioral development to enhance the treatment regimens for premature infants? 4) What exactly does a behavioral abnormality tell us about the CNS of the fetus?

Answers to these and other questions will greatly enrich our basic understanding of prenatal behavioral development and will demonstrate the clinical value of such scientifically based knowledge.

5.3 Outlook

One of the aims of this thesis was to investigate fetal behavior using novel methods with a particular focus on automated systems to make fine motion tracking of fetal body parts. The annotated dataset we generated and described in Chapter 3 through the help of a pose estimation algorithm is a starting point. The next step will be to perform a careful analysis of fetal upper limb kinematics. Once extracted reliable movement parameters, multiple research questions should be investigated.

A primary topic to be fully addressed in future studies is the onset of motor control. Findings reported by studies assessing the kinematics of fetal reaching movements should be replicated by collecting US scans at higher temporal resolution. Moreover, the traditional methodologies to quantify fetal movements should be replaced by automated pose estimation tools, such as DeepLabCut. The methodology to be used by these studies should be rethought and the results replicated within a quantitative and reliable framework. However, some of the most relevant questions on the existence of a prospective control in human fetuses remain crucial. Do fetuses already show anticipation (Delafield-Butt & Gangopadhyay, 2013)? Do fetuses prospectively calibrate their reaching patterning to the property of the target (Castiello et al., 2010; Parma et al., 2017; Zoia et al., 2013; Zoia et al., 2007)? New technologies will be required to massively serve the field of fetal psychology.

Studies examining anticipation in fetuses, as well as further replication of their results, could provide insight into neurodevelopmental disorders (e.g. ASD) in which the development of anticipations skills appears delayed (Sinha et al., 2014; Trevarthen & Delafield-Butt, 2013). Fetal movements in children who were later diagnosed with autism spectrum disorder (ASD) have not yet been studied. However, one could speculate that postnatal movement patterns of autistic children may be present in the womb (Gamliel 2012). Within this framework, the fetal movement's analysis stands as a promising tool to identify abnormal patterns of development in utero that are biomarkers of later onset of neurodevelopmental disorders (Castiello & Parma, 2018).

With the availability of fetal kinematic measures, one may wonder whether an individual style of movement exists before birth. Evidence showed that outputs of the motor system may vary substantially from one individual to another (Ting et al., 2015). Individual variations in movement patterns have been documented in a wide variety of animals (Calabrese et al., 2011), as well as in humans (Koul et al., 2016; Welch & Ting, 2008). Can we identify individual movement styles in human fetuses? Whether this is detected, is it consistent across different gestational periods (e.g., examining fetuses at around 13 weeks, then around 20 weeks of gestation)?

Assessing fetuses during the third trimester can raise some difficulties with measurement. Beyond eye, head, and body movements, other measures can be investigated in a fetus. For example, fetal heart rate (HR) and heart rate variability (HRV) are examined as an index of arousal (DiPietro et al., 2015). Prenatal measurements of HR and HRV can provide reliable and sensitive markers of fetal brain development that could potentially identify those at risk for stillbirth, sudden infant death syndrome (SIDS), or neurodevelopmental disorders (Shuffrey et al., 2019). Exploring options such as multiple measures (e.g., both fetal heart rate and movement or looking measurements) might thus provide additional pathways for a deeper understanding of fetal behavioral development.

Future studies employing a multidisciplinary approach are needed to further our knowledge of prenatal behavior and brain. Fetal functional neuroimaging research, acquiring data via optimized methods of magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI), has arisen as an

additional powerful approach for examining fetal brain function (Anderson & Thomason, 2013). Ultrasound and functional studies should make a threefold joint effort by i) optimizing research protocols to determine the best possible paradigms, ii) optimizing the stimulus to be delivered (if needed), iii) determining appropriate measures of fetal behavior (and how to interpret them). Methods to study human fetal development still need massive refinements.

Although the mystery is shrinking, fetal behavior remains alluring and promises many fundamental contributions to our science in the coming years.

References

- Abo-Yaqoub, S., Kurjak, A., Mohammed, A.-B., Shadad, A., & Abdel-Maaboud, M. (2012). The role of 4-D ultrasonography in prenatal assessment of fetal neurobehaviour and prediction of neurological outcome. *The Journal of Maternal-Fetal & Neonatal Medicine*, 25(3), 231-236. <https://doi.org/10.3109/14767058.2011.568552>
- AboEllail, M. A. M., Kanenishi, K., Mori, N., Noguchi, J., Marumo, G., & Hata, T. (2018). Ultrasound study of fetal movements in singleton and twin pregnancies at 12–19 weeks. *Journal of Perinatal Medicine*, 46(8), 832-838. <https://doi.org/10.1515/jpm-2017-0158>
- Adams, R. J., Courage, M. L., & Mercer, M. E. (1994). Systematic measurement of human neonatal color vision. *Vision Research*, 34(13), 1691-1701. [https://doi.org/10.1016/0042-6989\(94\)90127-9](https://doi.org/10.1016/0042-6989(94)90127-9)
- Anderson, A. L., & Thomason, M. E. (2013). Functional plasticity before the cradle: A review of neural functional imaging in the human fetus. *Neuroscience & Biobehavioral Reviews*, 37(9), 2220-2232. <https://doi.org/10.1016/j.neubiorev.2013.03.013>
- Andonotopo, W., & Kurjak, A. (2006). The assessment of fetal behavior of growth restricted fetuses by 4D sonography. *Journal of Perinatal Medicine*, 34(6), 471-478. <https://doi.org/10.1515/jpm.2006.092>
- Andonotopo, W., Medic, M., Salihagic-Kadic, A., Milenkovic, D., Maiz, N., & Scazzocchio, E. (2005). The assessment of fetal behavior in early pregnancy: comparison between 2D and 4D sonographic scanning. *Journal of Perinatal Medicine*, 33(5), 406-414. <https://doi.org/10.1515/jpm.2005.073>
- Ansuini, C., Cavallo, A., Koul, A., Jacono, M., Yang, Y., & Becchio, C. (2015). Predicting Object Size from Hand Kinematics: A Temporal Perspective. *PLoS ONE*, 10(3), e0120432. <https://doi.org/10.1371/journal.pone.0120432>
- Antsaklis, P., Porovic, S., Daskalakis, G., & Kurjak, A. (2017). 4D assessment of fetal brain function in diabetic patients. *Journal of Perinatal Medicine*, 45(6), 711-715. <https://doi.org/10.1515/jpm-2016-0394>
- Athanasiadis, A. P., Mikos, T., Tambakoudis, G. P., Theodoridis, T. D., Papastergiou, M., Assimakopoulos, E., & Tarlatzis, B. C. (2013). Neurodevelopmental fetal assessment using KANET scoring system in low and high risk pregnancies. *The Journal of Maternal-Fetal & Neonatal Medicine*, 26(4), 363-368. <https://doi.org/10.3109/14767058.2012.695824>

- Bardi, L., Regolin, L., & Simion, F. (2011). Biological motion preference in humans at birth: role of dynamic and configural properties. *Developmental Science*, 14(2), 353-359.
<https://doi.org/10.1111/j.1467-7687.2010.00985.x>
- Bardi, L., Regolin, L., & Simion, F. (2014). The first time ever I saw your feet: Inversion effect in newborns' sensitivity to biological motion. *Developmental Psychology*, 50(4), 986-993.
<https://doi.org/10.1037/a0034678>
- Barker, D. J. (1990). The fetal and infant origins of adult disease. *BMJ*, 301(6761), 1111.
<https://doi.org/10.1136/bmj.301.6761.1111>
- Barker, D. J. (2004). The developmental origins of chronic adult disease. *Acta Paediatrica*, 93, 26-33.
<https://doi.org/10.1080/08035320410022730>
- Barlow, J., Herath Ni Fau - Bartram Torrance, C., Bartram Torrance C Fau - Bennett, C., Bennett C Fau - Wei, Y., & Wei, Y. (2018). The Neonatal Behavioral Assessment Scale (NBAS) and Newborn Behavioral Observations (NBO) system for supporting caregivers and improving outcomes in caregivers and their infants. *Cochrane Database Syst Rev.*, 14(3).
<https://doi.org/10.1002/14651858.CD011754.pub2>
- Becchio, C., Koul, A., Ansuini, C., Bertone, C., & Cavallo, A. (2018). Seeing mental states: An experimental strategy for measuring the observability of other minds. *Physics of Life Reviews*, 24, 67-80. <https://doi.org/10.1016/j.plrev.2017.10.002>
- Bidet-Ildei, C., Kitromilides, E., Orliaguet, J.-P., Pavlova, M., & Gentaz, E. (2014). Preference for point-light human biological motion in newborns: Contribution of translational displacement. *Developmental Psychology*, 50(1), 113-120. <https://doi.org/10.1037/a0032956>
- Bot, M., Vladareanu, R., Burnei, A., Munteanu, A., Calo, I., & Vladareanu, S. (2020). Monochorionic vs Dichorionic Twins: Kanet Test vs Postnatal Neurodevelopment. *Maedica (Bucur)*, 15(1), 61-70.
<https://doi.org/10.26574/maedica.2020.15.1.61>
- Burr, D. C., Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually Inexperienced Chicks Exhibit Spontaneous Preference for Biological Motion Patterns. *PLoS Biology*, 3(7).
<https://doi.org/10.1371/journal.pbio.0030208>
- Calabrese, R. L., Norris, B. J., Wenning, A., & Wright, T. M. (2011). Coping with Variability in Small Neuronal Networks. *Integrative and Comparative Biology*, 51(6), 845-855.
<https://doi.org/10.1093/icb/icr074>

- Castiello, U., Becchio, C., Zoia, S., Nelini, C., Sartori, L., Blason, L., D'Ottavio, G., Bulgheroni, M., & Gallese, V. (2010). Wired to Be Social: The Ontogeny of Human Interaction. *PLoS ONE*, 5(10), e13199. <https://doi.org/10.1371/journal.pone.0013199>
- Castiello, U., & Parma, V. (2018). Fetal Kinematics: Basic Outcomes and Translational Outlook. *ACS Chemical Neuroscience*, 9(2), 165-166. <https://doi.org/10.1021/acschemneuro.8b00016>
- Chitty, L., & Altman, D. (2016). Charts of Fetal Size. *BMUS Bulletin*, 2(4), 9-19. <https://doi.org/10.1177/1742271x9400200404>
- Ciaunica, A., Safron, A., & Delafield-Butt, J. (2021). Back to square one: the bodily roots of conscious experiences in early life. *Neuroscience of Consciousness*, 7(2), 1-10. <https://doi.org/10.1093/nc/niab037>
- Craigheo, L., Lunghi, M., Leo, I., Ghirardi, V., & Simion, F. (2017). Newborns' attention is driven by the translational movement. *Visual Cognition*, 24(9-10), 487-498. <https://doi.org/10.1080/13506285.2017.1322651>
- Dang, Q., Yin, J., Wang, B., & Zheng, W. (2019). Deep learning based 2D human pose estimation: A survey. *Tsinghua Science and Technology*, 24(6), 663-676. <https://doi.org/10.26599/tst.2018.9010100>
- Das, R., Jana, N., Arora, N., & Sengupta, S. (2019). Ultrasound assessment of fetal hearing response to vibroacoustic stimulation. *The Journal of Maternal-Fetal & Neonatal Medicine*, 33(14), 2326-2332. <https://doi.org/10.1080/14767058.2018.1548600>
- de Klerk, C. C. J. M., Southgate, V., & Csibra, G. (2016). Predictive action tracking without motor experience in 8-month-old infants. *Brain and Cognition*, 109, 131-139. <https://doi.org/10.1016/j.bandc.2016.09.010>
- de Vries, J. I. P., & Fong, B. F. (2006). Normal fetal motility: an overview. *Ultrasound in Obstetrics and Gynecology*, 27(6), 701-711. <https://doi.org/10.1002/uog.2740>
- de Vries, J. I. P., Visser, G. H. A., & Prechtl, H. F. R. (1982). The emergence of fetal behaviour. I. Qualitative aspects. *Early Human Development*, 7(4), 301-322. [https://doi.org/10.1016/0378-3782\(82\)90033-0](https://doi.org/10.1016/0378-3782(82)90033-0)
- DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., & Maugeais, R. (1994, 1994/04/01/). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, 17(2), 159-164. [https://doi.org/https://doi.org/10.1016/0163-6383\(94\)90051-5](https://doi.org/https://doi.org/10.1016/0163-6383(94)90051-5)

- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9(2), 133-150. [https://doi.org/10.1016/0163-6383\(86\)90025-1](https://doi.org/10.1016/0163-6383(86)90025-1)
- Degani, S., Leibovitz, Z., Shapiro, I., & Ohel, G. (2009). Twins' temperament: early prenatal sonographic assessment and postnatal correlation. *Journal of Perinatology*, 29(5), 337-342. <https://doi.org/10.1038/jp.2008.238>
- Del Giudice, M. (2011). Alone in the dark? Modeling the conditions for visual experience in human fetuses. *Developmental Psychobiology*, 53(2), 214-219. <https://doi.org/10.1002/dev.20506>
- Delafield-Butt, J. T., & Gangopadhyay, N. (2013). Sensorimotor intentionality: The origins of intentionality in prospective agent action. *Developmental Review*, 33(4), 399-425. <https://doi.org/10.1016/j.dr.2013.09.001>
- Deng, J., Dong, W., Socher, R., Li, L.-J., Kai, L., & Li, F.-F. (2009). *ImageNet: A large-scale hierarchical image database* 2009 IEEE Conference on Computer Vision and Pattern Recognition,
- Di Giorgio, E., Frasnelli, E., Rosa Salva, O., Luisa Scattoni, M., Puopolo, M., Tosoni, D., Simion, F., & Vallortigara, G. (2016). Difference in Visual Social Predispositions Between Newborns at Low- and High-risk for Autism. *Scientific Reports*, 6(1). <https://doi.org/10.1038/srep26395>
- Dieb, A., Salam, R., Shaheen, D., & Shaeer, E. (2019). Evaluation of foetal neurological behaviour in hypothyroid pregnant females – a pilot study. *The Journal of Maternal-Fetal & Neonatal Medicine*, 32(16), 2617-2621. <https://doi.org/10.1080/14767058.2018.1442428>
- DiPietro, J. A. (2005). Neurobehavioral assessment before birth. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 4-13. <https://doi.org/10.1002/mrdd.20047>
- DiPietro, J. A., Costigan, K. A., & Voegtline, K. M. (2015). Studies in Fetal Behavior: Revisited, Renewed, and Reimagined. *Monographs of the Society for Research in Child Development*, 80(3), vii;1-94. <https://doi.org/10.1111/mono.v80.3>
- Donovan, T., Dunn, K., Penman, A., Young, R. J., & Reid, V. M. (2020). Fetal eye movements in response to a visual stimulus. *Brain and Behavior*, 10(8). <https://doi.org/10.1002/brb3.1676>
- Einspieler, C., Prechtl, H. F. R., Ferrari, F., Cioni, G., & Bos, A. F. (1997, 1997/11/24/). The qualitative assessment of general movements in preterm, term and young infants — review of the methodology. *Early Human Development*, 50(1), 47-60. [https://doi.org/https://doi.org/10.1016/S0378-3782\(97\)00092-3](https://doi.org/https://doi.org/10.1016/S0378-3782(97)00092-3)

- Eswaran, H., Lowery, C. L., Wilson, J. D., Murphy, P., & Preissl, H. (2004). Functional development of the visual system in human fetus using magnetoencephalography. *Experimental Neurology*, 190, 52-58. <https://doi.org/10.1016/j.expneurol.2004.04.007>
- Eswaran, H., Wilson J Fau - Preissl, H., Preissl H Fau - Robinson, S., Robinson S Fau - Vrba, J., Vrba J Fau - Murphy, P., Murphy P Fau - Rose, D., Rose D Fau - Lowery, C., & Lowery, C. (2002). Magnetoencephalographic recordings of visual evoked brain activity in the human fetus. *Lancet*(0140-6736), 779-780. [https://doi.org/10.1016/s0140-6736\(02\)09905-1](https://doi.org/10.1016/s0140-6736(02)09905-1)
- Eyre, J. A., Taylor, J. P., Villagra, F., Smith, M., & Miller, S. (2001). Evidence of activity-dependent withdrawal of corticospinal projections during human development. *Neurology*, 57(9), 1543-1554. <https://doi.org/10.1212/wnl.57.9.1543>
- Fagard, J., Esseily, R., Jacquy, L., O'Regan, K., & Somogyi, E. (2018). Fetal Origin of Sensorimotor Behavior. *Frontiers in Neurorobotics*, 12(23), 1-7. <https://doi.org/10.3389/fnbot.2018.00023>
- Ferrari, G. A., Nicolini, Y., Demuru, E., Tosato, C., Hussain, M., Scesa, E., Romei, L., Boerci, M., Iappini, E., Dalla Rosa Prati, G., Palagi, E., & Ferrari, P. F. (2016). Ultrasonographic Investigation of Human Fetus Responses to Maternal Communicative and Non-communicative Stimuli. *Frontiers in Psychology*, 7(354), 1-10. <https://doi.org/10.3389/fpsyg.2016.00354>
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7), 1688-1703. <https://doi.org/10.1523/jneurosci.05-07-01688.1985>
- Fulford, J., Vadeyar, S. H., Dodampahala, S. H., Moore, R. J., Young, P., Baker, P. N., James, D. K., & Gowland, P. A. (2003). Fetal brain activity in response to a visual stimulus. *Human Brain Mapping*, 20(4), 239-245. <https://doi.org/10.1002/hbm.10139>
- Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *The Journal of Neuroscience*, 2(11), 1527-1537. <https://doi.org/10.1523/jneurosci.02-11-01527.1982>
- Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., Taxidis, J., Najafi, F., Gauthier, J. L., Zhou, P., Khakh, B. S., Tank, D. W., Chklovskii, D. B., & Pnevmatikakis, E. A. (2019). CalmAn an open source tool for scalable calcium imaging data analysis. *eLife*, 8. <https://doi.org/10.7554/eLife.38173>
- Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces Attract Infants' Attention in Complex Displays. *Infancy*, 14(5), 550-562. <https://doi.org/10.1080/15250000903144199>

- Green, S., & Sparling, J. W. (1993). Q-MOVE: Development of a qualitative assessment of fetal movement. *Physical & Occupational Therapy In Pediatrics*, 12(2-3), 115-137.
- Hadad, B.-S., Maurer, D., & Lewis, T. L. (2012). Sparing of sensitivity to biological motion but not of global motion after early visual deprivation. *Developmental Science*, 15(4), 474-481. <https://doi.org/10.1111/j.1467-7687.2012.01145.x>
- Hadders-Algra, M. (2018). Early human motor development: From variation to the ability to vary and adapt. *Neuroscience & Biobehavioral Reviews*, 90, 411-427. <https://doi.org/10.1016/j.neubiorev.2018.05.009>
- Haggard, P., & Wolpert, D. M. (2005). Disorders of Body Scheme. In *High-order motor disorders: from neuroanatomy and neurobiology to clinical neurology* (pp. 261–271). Oxford, UK: Oxford University Press.
- Hagihara, Y., Nakajima, A., Fukuda, S., Goto, S., Iida, H., & Yamazaki, M. (2009). Running Exercise for Short Duration Increases Bone Mineral Density of Loaded Long Bones in Young Growing Rats. *The Tohoku Journal of Experimental Medicine*, 219(2), 139-143. <https://doi.org/10.1620/tjem.219.139>
- Hakamada, S., Hayakawa, F., Kuno, K., & Tanaka, R. (1988). Development of the monosynaptic reflex pathway in the human spinal cord. *Developmental Brain Research*, 42(2), 239-246. [https://doi.org/10.1016/0165-3806\(88\)90242-8](https://doi.org/10.1016/0165-3806(88)90242-8)
- Hanaoka, U., Hata, T., Kanenishi, K., Mostafa AboEllail, M. A., Uematsu, R., Konishi, Y., Kusaka, T., Noguchi, J., Marumo, G., Vasilj, O., & Kurjak, A. (2016). Does ethnicity have an effect on fetal behavior? A comparison of Asian and Caucasian populations. *Journal of Perinatal Medicine*, 44(2), 217-221. <https://doi.org/10.1515/jpm-2015-0036>
- Hata, T., Hanaoka, U., AboEllail, M. A. M., Uematsu, R., Noguchi, J., Kusaka, T., & Kurjak, A. (2016a). Is there a sex difference in fetal behavior? A comparison of the KANET test between male and female fetuses. *Journal of Perinatal Medicine*, 44(5), 585-588. <https://doi.org/10.1515/jpm-2015-0387>
- Hata, T., Hanaoka, U., AboEllail, M. A. M., Uematsu, R., Noguchi, J., Kusaka, T., & Kurjak, A. (2016b). Does Parity have an Effect on Fetal Behavior? A Comparison between Primi- and Multiparas. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 10(1), 99-102. <https://doi.org/10.5005/jp-journals-10009-1448>
- Hata, T., Kanenishi, K., Mori, N., AboEllail, M. A. M., Hanaoka, U., Koyano, K., Kato, I., & Kusaka, T. (2019). Prediction of postnatal developmental disabilities using the antenatal fetal

- neurodevelopmental test: KANET assessment. *Journal of Perinatal Medicine*, 47(1), 77-81. <https://doi.org/10.1515/jpm-2018-0169>
- Hata, T., Kanenishi, K., Mori, N., AboEllail, M. A. M., Hanaoka, U., Koyano, K., Kato, I., Kusaka, T., & Kurjak, A. (2019). Mini KANET: Simple Fetal Antenatal Neurodevelopmental Test. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 13(2), 59-63. <https://doi.org/10.5005/jp-journals-10009-1586>
- Hata, T., Kanenishi, K., & Sasaki, M. (2010). Four-dimensional sonographic assessment of fetal movement in the late first trimester. *International Journal of Gynecology & Obstetrics*, 109(3), 190-193. <https://doi.org/10.1016/j.ijgo.2009.12.020>
- Hata, T., Kanenishi, K., Sasaki, M., & Yanagihara, T. (2011). Fetal reflex movement in twin pregnancies late in the first trimester: 4-D sonographic study. *Ultrasound in Medicine & Biology*, 37(11), 1948-1951. <https://doi.org/10.1016/j.ultrasmedbio.2011.06.005>
- Hata, T., Sasaki, M., & Yanagihara, T. (2012). Difference in the frequency of types of inter-twin contact at 10–13 weeks' gestation: preliminary four-dimensional sonographic study. *The Journal of Maternal-Fetal & Neonatal Medicine*, 25(3), 226-230. <https://doi.org/10.3109/14767058.2011.568551>
- Hendrickson, A., Possin, D., Vajzovic, L., & Toth, C. A. (2012). Histologic Development of the Human Fovea From Midgestation to Maturity. *American Journal of Ophthalmology*, 154(5), 767-778.e762. <https://doi.org/10.1016/j.ajo.2012.05.007>
- Hepper, P. G. (1996). Fetal behavior: why so sceptical? *Ultrasound in Obstetrics & Gynecology*, 8, 145-148.
- Hepper, P. G., Scott, D., & Shahidullah, S. (1993). Newborn and fetal response to maternal voice. *Journal of Reproductive and Infant Psychology*, 11(3), 147-153. <https://doi.org/10.1080/02646839308403210>
- Hielkema, T., & Hadders-Algra, M. (2016). Motor and cognitive outcome after specific early lesions of the brain - a systematic review. *Developmental Medicine & Child Neurology*, 58, 46-52. <https://doi.org/10.1111/dmcn.13047>
- Hofsten, C. V. (2009). Action, the foundation for cognitive development. *Scandinavian Journal of Psychology*, 50(6), 617-623. <https://doi.org/10.1111/j.1467-9450.2009.00780.x>
- Honemeyer, U., Talic, A., Therwat, A., Paulose, L., & Patidar, R. (2013). The clinical value of KANET in studying fetal neurobehavior in normal and at-risk pregnancies. *Journal of Perinatal Medicine*, 41(2), 187-197. <https://doi.org/10.1515/jpm-2011-0251>

- Hooker, D. (1960). Development reaction to environment. *Yale J Biol Med.*, 32(6), 431-440.
- Hopkins, B., & Johnson, S. P. (2005). *Prenatal Development of Postnatal Functions*. Praeger Publishers.
<https://books.google.com.sv/books?id=Ylp7a8ZJxe4C>
- Inoue, M. F., Koyanagi, T. F., Nakahara, H. F., Hara, K., Hori, E., & Nakano, H. (1986). Functional development of human eye movement in utero assessed quantitatively with real-time ultrasound. *American Journal of Obstetrics and Gynecology*, 155(1)(0002-9378), 170-174.
- Johansson, G. (1973, 1973/06/01). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201-211. <https://doi.org/10.3758/BF03212378>
- Joseph, R. (2000). Fetal Brain Behavior and Cognitive Development. *Developmental Review*, 20(1), 81-98. <https://doi.org/10.1006/drev.1999.0486>
- Kim, T. H., Lee, J. J., Chung, S. H., Lee, H. H., Lee, K. H., Choi, K. Y., & Lee, S. H. (2010). Efficacy of assessment in fetal behaviour by four dimensional ultrasonography. *Journal of Obstetrics and Gynaecology*, 30(5), 439-443. <https://doi.org/10.3109/01443615.2010.486084>
- Kisilevsky, B. S., Hains Sm Fau - Lee, K., Lee K Fau - Xie, X., Xie X Fau - Huang, H., Huang H Fau - Ye, H. H., Ye Hh Fau - Zhang, K., Zhang K Fau - Wang, Z., & Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychol. Sci*, 14(0956-7976 (Print)), 220-224. <https://doi.org/10.1111/1467-9280.02435>
- Kisilevsky, B. S., & Low, J. A. (1998). Human Fetal Behavior: 100 Years of Study. *Developmental Review*, 18(1), 1-29. <https://doi.org/10.1006/drev.1998.0452>
- Kiuchi, M., Nagata, N., Ikeno, S., & Terakawa, N. (2000). The relationship between the response to external light stimulation and behavioral states in the human fetus: how it differs from vibroacoustic stimulation. *Early Human Development*, 58(2), 153-165.
[https://doi.org/10.1016/s0378-3782\(00\)00074-8](https://doi.org/10.1016/s0378-3782(00)00074-8)
- Klin, A., & Jones, W. (2008). Altered face scanning and impaired recognition of biological motion in a 15-month-old infant with autism. *Developmental Science*, 11(1), 40-46.
<https://doi.org/10.1111/j.1467-7687.2007.00608.x>
- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459(7244), 257-261.
<https://doi.org/10.1038/nature07868>

- Koul, A., Cavallo, A., Ansuini, C., & Becchio, C. (2016). Doing It Your Way: How Individual Movement Styles Affect Action Prediction. *PLoS ONE*, 11(10), 1-14.
<https://doi.org/10.1371/journal.pone.0165297>
- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, 93(3), 480-490.
<https://doi.org/10.1016/j.neuron.2016.12.041>
- Kuniyoshi, Y. (2019). Fusing autonomy and sociality via embodied emergence and development of behaviour and cognition from fetal period. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1771). <https://doi.org/10.1098/rstb.2018.0031>
- Kuno, A., Akiyama, M., Yamashiro, C., Tanaka, H., Yanagihara, T., & Hata, T. (2001). Three-dimensional sonographic assessment of fetal behavior in the early second trimester of pregnancy. *Journal of Ultrasound in Medicine*, 20(12), 1271-1275. <https://doi.org/10.7863/jum.2001.20.12.1271>
- Kurjak, A., Abo-Yaqoub, S., Stanojevic, M., Yigiter, A. B., Vasilj, O., Lebit, D., Shaddad, A. N., Ahmed, B., Kavak, Z. N., Miskovic, B., Vladareanu, R., Barisic, L. S., Azumendi, G., Younis, M., Pooh, R. K., & Kadic, A. S. (2010). The potential of 4D sonography in the assessment of fetal neurobehavior – multicentric study in high-risk pregnancies. *Journal of Perinatal Medicine*, 38(1), 77-82.
<https://doi.org/10.1515/jpm.2010.012>
- Kurjak, A., Andonotopo, W., Hafner, T., Salihagic Kadic, A., Stanojevic, M., Azumendi, G., Ahmed, B., Carrera, J. M., & Troyano, J. M. (2006). Normal standards for fetal neurobehavioral developments – longitudinal quantification by four-dimensional sonography. *Journal of Perinatal Medicine*, 34(1), 56-65. <https://doi.org/10.1515/jpm.2006.007>
- Kurjak, A., Antsaklis, P., Stanojevic, M., & Porovic, S. (2017). Fetal Behavior assessed by Four-dimensional Sonography. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 11(2), 146-168. <https://doi.org/10.5005/jp-journals-10009-1516>
- Kurjak, A., Antsaklis, P., Stanojevic, M., Vladareanu, R., Vladareanu, S., Neto, R. M., Barisic, L. S., Porovic, S., & Delic, T. (2017). Multicentric studies of the fetal neurobehavior by KANET test. *Journal of Perinatal Medicine*, 45(6), 717-727. <https://doi.org/10.1515/jpm-2016-0409>
- Kurjak, A., Azumendi, G., Veček, N., Kupešić, S., Solak, M., Varga, D., & Chervenak, F. (2003). Fetal hand movements and facial expression in normal pregnancy studied by four-dimensional sonography. *Journal of Perinatal Medicine*, 31(6), 496-508.
<https://doi.org/10.1515/jpm.2003.076>
- Kurjak, A., Miskovic, B., Stanojevic, M., Amiel-Tison, C., Ahmed, B., Azumendi, G., Vasilj, O., Andonotopo, W., Turudic, T., & Salihagic-Kadic, A. (2008). New scoring system for fetal

- neurobehavior assessed by three- and four-dimensional sonography. *Journal of Perinatal Medicine*, 36(1), 73-81. <https://doi.org/10.1515/jpm.2008.007>
- Kurjak, A., Stanojevic, M., Andonotopo, W., Salihagic-Kadic, A., Carrera, J. M., & Azumendi, G. (2004). Behavioral pattern continuity from prenatal to postnatal life a study by four-dimensional (4D) ultrasonography. *Journal of Perinatal Medicine*, 32(4), 346-353. <https://doi.org/10.1515/jpm.2004.065>
- Kurjak, A., Stanojevic, M., Andonotopo, W., Scazzocchio-Duenas, E., Azumendi, G., & Carrera, J. M. (2005). Fetal behavior assessed in all three trimesters of normal pregnancy by four-dimensional ultrasonography. *Croatian medical journal*, 46(5), 772-780.
- Kurjak, A., Talic, A., Honemeyer, U., Stanojevic, M., & Zalud, I. (2013). Comparison between antenatal neurodevelopmental test and fetal Doppler in the assessment of fetal well being. *Journal of Perinatal Medicine*, 41(1), 107-114. <https://doi.org/10.1515/jpm-2012-0018>
- Kurjak, A., Talic, A., Stanojevic, M., Honemeyer, U., Serra, B., Prats, P., & Di Renzo, G. C. (2013). The study of fetal neurobehavior in twins in all three trimesters of pregnancy. *The Journal of Maternal-Fetal & Neonatal Medicine*, 26(12), 1186-1195. <https://doi.org/10.3109/14767058.2013.773306>
- Kurjak, A., Vecek, N., Hafner, T., Bozek, T., Funduk-Kurjak, B., & Ujevic, B. (2002). Prenatal diagnosis: what does four-dimensional ultrasound add? *Journal of Perinatal Medicine*, 30(1), 57-62. <https://doi.org/10.1515/jpm.2002.008>
- Landis, J. R., & Koch, G. G. (1977, Mar). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://www.ncbi.nlm.nih.gov/pubmed/843571>
- Lebit, F.-D., & Vladareanu, R. (2011). The Role of 4D Ultrasound in the Assessment of Fetal Behaviour. *Maedica (Bucur)*, 6(2), 120-127.
- Lecanuet, J.-P., Fifer, W. P., Krasnegor, N. A., & Smotherman, W. P. (2013). *Fetal development: A psychobiological perspective*. Psychology Press.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, 62(10), e1-34. <https://doi.org/10.1016/j.jclinepi.2009.06.006>
- Manning, F. A., Harman, C. R., Morrison, I., Menticoglou, S. M., Lange, I. R., & Johnson, J. M. (1990, 1990/03/01/). Fetal assessment based on fetal biophysical profile scoring: IV. An analysis of

- perinatal morbidity and mortality. *American Journal of Obstetrics and Gynecology*, 162(3), 703-709. [https://doi.org/https://doi.org/10.1016/0002-9378\(90\)90990-O](https://doi.org/https://doi.org/10.1016/0002-9378(90)90990-O)
- Marx, V., & Nagy, E. (2015). Fetal Behavioural Responses to Maternal Voice and Touch. *PLoS ONE*, 10(6), e0129118. <https://doi.org/10.1371/journal.pone.0129118>
- Marx, V., & Nagy, E. (2017). Fetal behavioral responses to the touch of the mother's abdomen: A Frame-by-frame analysis. *Infant Behavior and Development*, 47, 83-91. <https://doi.org/10.1016/j.infbeh.2017.03.005>
- Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci*, 21, 1281-1289. <https://doi.org/10.1038/s41593-018-0209-y>
- Mathis, A., Schneider, S., Lauer, J., & Mathis, M. W. (2020). A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. *Neuron*, 108(1), 44-65. <https://doi.org/10.1016/j.neuron.2020.09.017>
- Matuz, T., Govindan, R. B., Preissl, H., Siegel, E. R., Muenssinger, J., Murphy, P., Ware, M., Lowery, C. L., & Eswaran, H. (2012). Habituation of visual evoked responses in neonates and fetuses: A MEG study. *Developmental Cognitive Neuroscience*, 2(3), 303-316. <https://doi.org/10.1016/j.dcn.2012.03.001>
- McCubbin, J., Murphy, P., Eswaran, H., Preissl, H., Yee, T., Robinson, S. E., & Vrba, J. (2007). Validation of the flash-evoked response from fetal MEG. *Physics in Medicine and Biology*, 52(19), 5803-5813. <https://doi.org/10.1088/0031-9155/52/19/005>
- Merz, E. (2005). Current 3D/4D ultrasound technology in prenatal diagnosis. *European Clinics in Obstetrics and Gynaecology*, 1(3), 184-193. <https://doi.org/10.1007/s11296-005-0005-6>
- Miskovic, B., Vasilj, O., Stanojevic, M., Ivanković, D., Kerner, M., & Tikvica, A. (2010). The comparison of fetal behavior in high risk and normal pregnancies assessed by four dimensional ultrasound. *The Journal of Maternal-Fetal & Neonatal Medicine*, 23(12), 1461-1467. <https://doi.org/10.3109/14767051003678200>
- Moran, D. W., & Schwartz, A. B. (1999). Motor cortical representation of speed and direction during reaching. *Journal of Neurophysiology*, 82(5), 2676-2692. <https://doi.org/10.1152/jn.1999.82.5.2676>
- Morante, A., Dubowitz, L. M. S., Levene, M., & Dubowitz, V. (2008). The Development of Visual Function in Normal and Neurologically Abnormal Preterm and Fullterm Infants. *Developmental Medicine & Child Neurology*, 24, 771-784. <https://doi.org/10.1111/j.1469-8749.1982.tb13698.x>

- Mori, N., Kanenishi, K., AboEllail, M. A. M., Nitta, E., Noguchi, J., Marumo, G., & Hata, T. (2018). Singleton and Twin Fetal Movements before 20 Weeks of Gestation. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 12(2), 99-103. <https://doi.org/10.5005/jp-journals-10009-1558>
- Moyer-Mileur, L. J., Ball, S. D., Brunstetter, V. L., & Chan, G. M. (2008). Maternal-administered physical activity enhances bone mineral acquisition in premature very low birth weight infants. *Journal of Perinatology*, 28(6), 432-437. <https://doi.org/10.1038/jp.2008.17>
- Mulder, E. J. H., & Visser, G. H. A. (2016). Fetal Behavior: Clinical and Experimental Research in the Human. In *Fetal Development* (pp. 87-105). https://doi.org/10.1007/978-3-319-22023-9_5
- Mündermann, L., Corazza, S., & Andriacchi, T. P. (2006). The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *Journal of NeuroEngineering and Rehabilitation*, 3(6). <https://doi.org/10.1186/1743-0003-3-6>
- Myowa-Yamakoshi, M., & Takeshita, H. (2006). Do Human Fetuses Anticipate Self-Oriented Actions? A Study by Four-Dimensional (4D) Ultrasonography. *Infancy*, 10(3), 289-301. https://doi.org/10.1207/s15327078in1003_5
- Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols*, 14(7), 2152-2176. <https://doi.org/10.1038/s41596-019-0176-0>
- Neto, R. M. (2015). KANET in Brazil: First Experience. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 9(1), 1-5. <https://doi.org/10.5005/jp-journals-10009-1384>
- Nijhuis, J. G., Prechtl, H. F. R., Martin, C. B., & Bots, R. S. G. M. (1982). Are there behavioural states in the human fetus? *Early Human Development*, 6(2), 177-195. [https://doi.org/10.1016/0378-3782\(82\)90106-2](https://doi.org/10.1016/0378-3782(82)90106-2)
- Nyquist, H. (1928). Certain Topics in Telegraph Transmission Theory. *Transactions of the American Institute of Electrical Engineers*, 47(2), 617-644. <https://doi.org/10.1109/T-AIEE.1928.5055024>
- Oakes, L. M. (2010). Using Habituation of Looking Time to Assess Mental Processes in Infancy. *Journal of Cognition and Development*, 11(3), 255-268. <https://doi.org/10.1080/15248371003699977>
- Ogo, K., Kanenishi, K., Mori, N., AboEllail, M. A. M., & Hata, T. (2019). Change in fetal behavior in response to vibroacoustic stimulation. *Journal of Perinatal Medicine*, 47(5), 558-563. <https://doi.org/10.1515/jpm-2018-0344>

- Okado, N. (1981). Onset of synapse formation in the human spinal cord. *J Comp Neurol.*, 201(2), 211-219. <https://doi.org/10.1002/cne.902010206>
- Okado, N., & Kojima, T. (1984). *Ontogeny of the central nervous system: neurogenesis, fibre connection, synaptogenesis and myelination in the spinal cord*. Oxford: Blackwell Science.
- Okawa, H., Morokuma, S., Maehara, K., Arata, A., Ohmura, Y., Horinouchi, T., Konishi, Y., & Kato, K. (2017). Eye movement activity in normal human fetuses between 24 and 39 weeks of gestation. *PLoS ONE*, 12(7), 1-12. <https://doi.org/10.1371/journal.pone.0178722>
- Parma, V., Brasselet, R., Zoia, S., Bulgheroni, M., & Castiello, U. (2017). The origin of human handedness and its role in pre-birth motor control. *Scientific Reports*, 7(16804). <https://doi.org/10.1038/s41598-017-16827-y>
- Pavlova, M. A. (2011). Biological Motion Processing as a Hallmark of Social Cognition. *Cerebral Cortex*, 22(5), 981-995. <https://doi.org/10.1093/cercor/bhr156>
- Pereira, T. D., Shaevitz, J. W., & Murthy, M. (2020). Quantifying behavior to understand the brain. *Nature Neuroscience*, 23(12), 1537-1549. <https://doi.org/10.1038/s41593-020-00734-z>
- Piontelli, A. (2010). *Development of Normal Fetal Movements. The First 25 Weeks of Gestation Introduction*. Springer Verlag Italia. <https://doi.org/10.1007/978-88-470-1402-2>
- Piontelli, A. (2015). *Development of Normal Fetal Movements. The Last 15 Weeks of Gestation*. Springer Verlag Italia <https://doi.org/10.1007/978-88-470-5373-1>
- Prats, P., Serra, B., Fournier, S., Baulies, S., Andonotopo, W., & Kurjak, A. (2012). 4D Sonographic Assessment of Intertwin Contacts. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 6(2), 154-159. <https://doi.org/10.5005/jp-journals-10009-1237>
- Prechtl, H. F. R., & Einspieler, C. (1997). Is neurological assessment of the fetus possible? *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 75(1), 81-84. [https://doi.org/10.1016/s0301-2115\(97\)00197-8](https://doi.org/10.1016/s0301-2115(97)00197-8)
- Predojević, M., Talić, A., Stanojević, M., Kurjak, A., & Salihagić Kadić, A. (2014). Assessment of motoric and hemodynamic parameters in growth restricted fetuses – case study. *The Journal of Maternal-Fetal & Neonatal Medicine*, 27(3), 247-251. <https://doi.org/10.3109/14767058.2013.807241>

- Pretorius, D. H., & Nelson, T. R. (1995). Three-dimensional ultrasound. *Ultrasound in Obstetrics and Gynecology* 5, 219-221.
- Rao, S., Chun, C., Fan, J., Kofron, J. M., Yang, M. B., Hegde, R. S., Ferrara, N., Copenhagen, D. R., & Lang, R. A. (2013). A direct and melanopsin-dependent fetal light response regulates mouse eye development. *Nature*, 494(7436), 243-246. <https://doi.org/10.1038/nature11823>
- Reid, V. M., & Dunn, K. (2021). The Fetal Origins of Human Psychological Development. *Current Directions in Psychological Science*, 30(2), 144-150. <https://doi.org/10.1177/0963721420984419>
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Current Biology*, 27(12), 1825-1828. <https://doi.org/10.1016/j.cub.2017.05.044>
- Reissland, N., Aydin, E., Francis, B., & Exley, K. (2015). Laterality of foetal self-touch in relation to maternal stress. *Laterality: Asymmetries of Body, Brain and Cognition*, 20(1), 82-94. <https://doi.org/10.1080/1357650x.2014.920339>
- Reissland, N., Francis, B., Aydin, E., Mason, J., & Schaal, B. (2014). The development of anticipation in the fetus: A longitudinal account of human fetal mouth movements in reaction to and anticipation of touch. *Developmental Psychobiology*, 56(5), 955-963. <https://doi.org/10.1002/dev.21172>
- Reissland, N., Francis, B., Buttanshaw, L., Austen, J. M., & Reid, V. (2016). Do fetuses move their lips to the sound that they hear? An observational feasibility study on auditory stimulation in the womb. *Pilot and Feasibility Studies*, 2(14), 1-7. <https://doi.org/10.1186/s40814-016-0053-3>
- Reissland, N., Francis, B., Kumarendran, K., & Mason, J. (2015). Ultrasound observations of subtle movements: a pilot study comparing foetuses of smoking and nonsmoking mothers. *Acta Paediatrica*, 104(6), 596-603. <https://doi.org/10.1111/apa.13001>
- Reissland, N., & Hopkins, B. (2010). Introduction: towards a fetal psychology. *Infant and Child Development*, 19(1), 1-5. <https://doi.org/10.1002/icd.662>
- Reissland, N., & Kisilevsky, B. S. (2016). *Fetal Development. Research on Brain and Behavior, Environmental Influences, and Emerging Technologies*. Springer International Publishing Switzerland 2016. <https://doi.org/10.1007/978-3-319-22023-9>
- Reissland, N., Wood, R., Einbeck, J., & Lane, A. (2020). Effects of maternal mental health on fetal visual preference for face-like compared to non-face like light stimulation. *Early Human Development*, 151, 105227. <https://doi.org/10.1016/j.earlhumdev.2020.105227>

- Roby-Brami, A., Bennis, N., Mokhtari, M., & Baraduc, P. (2000). Hand orientation for grasping depends on the direction of the reaching movement. *Brain Research*, 869(1-2), 121-129.
[https://doi.org/10.1016/s0006-8993\(00\)02378-7](https://doi.org/10.1016/s0006-8993(00)02378-7)
- Romeo, D. M., Ricci, D., Serrao, F., Gallini, F., Olivieri, G., Cota, F., Romagnoli, C., & Mercuri, E. (2012, May). Visual function assessment in late-preterm newborns. *Early Human Development*, 88(5), 301-305. <https://doi.org/10.1016/j.earlhumdev.2011.08.024>
- Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., & van der Wel, R. (2012). Cognition, action, and object manipulation. *Psychological Bulletin*, 138(5), 924-946.
<https://doi.org/10.1037/a0027839>
- Sajapala, S., AboEllail, M. A. M., Kanenishi, K., Mori, N., Marumo, G., & Hata, T. (2017). 4D ultrasound study of fetal movement early in the second trimester of pregnancy. *Journal of Perinatal Medicine*, 45(6), 737-743. <https://doi.org/10.1515/jpm-2016-0250>
- Salisbury, A. L., Fallone, M. D., & Lester, B. (2005). Neurobehavioral assessment from fetus to infant: The NICU network neurobehavioral scale and the fetal neurobehavior coding scale. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 14-20.
<https://doi.org/10.1002/mrdd.20058>
- Sarnat, H. (2003). Functions of the corticospinal and corticobulbar tracts in the human newborn. *Journal of Pediatric Neurology*, 01(01), 003-008. <https://doi.org/10.1055/s-0035-1557162>
- Sasaki, M., Yanagihara, T., Naitoh, N., & Hata, T. (2010). Four-dimensional sonographic assessment of inter-twin contact late in the first trimester. *International Journal of Gynecology & Obstetrics*, 108, 104-107. <https://doi.org/10.1016/j.ijgo.2009.09.025>
- Schaal, B. (2000). Human Foetuses Learn Odours from their Pregnant Mother's Diet. *Chemical Senses*, 25(6), 729-737. <https://doi.org/10.1093/chemse/25.6.729>
- Scott, S. H., Gribble, P. L., Graham, K. M., & Cabel, D. W. (2001). Dissociation between hand motion and population vectors from neural activity in motor cortex. *Nature*, 413, 161-165.
<https://doi.org/10.1038/35093102>
- Seckl, J. R., & Holmes, M. C. (2007). Mechanisms of disease: glucocorticoids, their placental metabolism and fetal 'programming' of adult pathophysiology. *Nat Clin Pract Endocrinol Metab*, 3(6), 479-488. <https://doi.org/10.1038/ncpendmet0515>
- Shannon, C. E. (1949). Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1), 10-21.
<https://doi.org/10.1109/jrproc.1949.232969>

- Sheridan, C. J., Preissl, H., Siegel, E. R., Murphy, P., Ware, M., Lowery, C. L., & Eswaran, H. (2008). Neonatal and fetal response decrement of evoked responses: A MEG study. *Clinical Neurophysiology*, 119(4), 796-804. <https://doi.org/10.1016/j.clinph.2007.11.174>
- Shuffrey, L. C., Myers, M. M., Odendaal, H. J., Elliott, A. J., du Plessis, C., Groenewald, C., Burd, L., Angal, J., Nugent, J. D., Isler, J. R., & Fifer, W. P. (2019). Fetal heart rate, heart rate variability, and heart rate/movement coupling in the Safe Passage Study. *Journal of Perinatology*, 39(5), 608-618. <https://doi.org/10.1038/s41372-019-0342-9>
- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences*, 105(2), 809-813. <https://doi.org/10.1073/pnas.0707021105>
- Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences*, 111(42), 15220-15225. <https://doi.org/10.1073/pnas.1416797111>
- Sival, D. A. (1993, 1993/09/01/). Studies on fetal motor behaviour in normal and complicated pregnancies. *Early Human Development*, 34(1), 13-20. [https://doi.org/https://doi.org/10.1016/0378-3782\(93\)90036-T](https://doi.org/https://doi.org/10.1016/0378-3782(93)90036-T)
- Talic, A., Kurjak, A., Ahmed, B., Stanojevic, M., Predojevic, M., Kadic, A. S., & Di Renzo, G. C. (2011). The potential of 4D sonography in the assessment of fetal behavior in high-risk pregnancies. *The Journal of Maternal-Fetal & Neonatal Medicine*, 24(7), 948-954. <https://doi.org/10.3109/14767058.2010.534830>
- Talic, A., Kurjak, A., Stanojevic, M., Honemeyer, U., Badreldeen, A., & DiRenzo, G. C. (2012). The assessment of fetal brain function in fetuses with ventrikulomegaly: the role of the KANET test. *The Journal of Maternal-Fetal & Neonatal Medicine*, 25(8), 1267-1272. <https://doi.org/10.3109/14767058.2011.634463>
- Tan, K. H., Smyth, R. M. D., & Wei, X. (2013). Fetal vibroacoustic stimulation for facilitation of tests of fetal wellbeing. *Cochrane Database of Systematic Reviews*(12). <https://doi.org/10.1002/14651858.CD002963.pub2>
- Ten Hof, J., Nijhuis, I. J. M., Nijhuis, J. G., Narayan, H., Taylor, D. J., Visser, G. H. A., & Mulder, E. J. H. (1999). Quantitative analysis of fetal general movements: methodological considerations. *Early Human Development*, 56(1), 57-73. [https://doi.org/10.1016/s0378-3782\(99\)00035-3](https://doi.org/10.1016/s0378-3782(99)00035-3)
- Ting, Lena H., Chiel, Hillel J., Trumbower, Randy D., Allen, Jessica L., McKay, J. L., Hackney, Madeleine E., & Kesar, Trisha M. (2015). Neuromechanical Principles Underlying Movement Modularity and

Their Implications for Rehabilitation. *Neuron*, 86(1), 38-54.
<https://doi.org/10.1016/j.neuron.2015.02.042>

Tinjić, S. (2019). Experiences and Results of the KANET Test Application in Clinical Practice in Tuzla, Bosnia and Herzegovina. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 13(3), 94-98. <https://doi.org/10.5005/jp-journals-10009-1595>

Trevarthen, C., & Delafield-Butt, J. T. (2013). Autism as a developmental disorder in intentional movement and affective engagement. *Frontiers in Integrative Neuroscience*, 7, 1-16.
<https://doi.org/10.3389/fnint.2013.00049>

Unver Dogan, N., Uysal, I. I., Karabulut, A. K., & Fazliogullari, Z. (2009). The motor branches of median and ulnar nerves that innervate superficial flexor muscles: a study in human fetuses. *Surgical and Radiologic Anatomy*, 32(3), 225-233. <https://doi.org/10.1007/s00276-009-0580-1>

Vallortigara, G., & Regolin, L. (2006). Gravity bias in the interpretation of biological motion by inexperienced chicks. *Current Biology*, 16(8), R279-R280.
<https://doi.org/10.1016/j.cub.2006.03.052>

Vindla, S., & James, D. (1995). Fetal behaviour as a test of fetal wellbeing. *BJOG: An International Journal of Obstetrics and Gynaecology*, 102(8), 597-599. <https://doi.org/10.1111/j.1471-0528.1995.tb11395.x>

Visser, G. H., F., M. E., & Prechtl, H. F. (1992). Studies on developmental neurology in the human fetus. *Developmental Pharmacology and Therapeutics*, 18(3-4)(0379-8305), 175-183.

Vladareanu, R., Lebit, D., & Constantinescu, S. (2012). Ultrasound assessment of fetal neurobehaviour in high-risk pregnancies. *Donald School J Ultrasound Obstetrics Gynecology*, 6(2), 132-147.
<https://doi.org/10.5005/jp-journals-10009-1235>

Watterberg, K. L., Aucott, S., Benitz, W. E., Cummings, J. J., Eichenwald, E. C., Goldsmith, J., Poindexter, B. B., Puopolo, K., Stewart, D. L., Wang, K. S., Ecker, J. L., Wax, J. R., Borders, A. E. B., El-Sayed, Y. Y., Heine, R. P., Jamieson, D. J., Mascola, M. A., Minkoff, H. L., Stuebe, A. M., Sumners, J. E., Tuuli, M. G., & Wharton, K. R. (2015). The Apgar Score. *Pediatrics*, 136(4), 819-822.
<https://doi.org/10.1542/peds.2015-2651>

Welch, T. D., & Ting, L. H. (2008, Feb). A feedback model reproduces muscle activity during human postural responses to support-surface translations. *Journal of Neurophysiology*, 99(2), 1032-1038. <https://doi.org/10.1152/jn.01110.2007>

- World Medical, A. (2013, Nov 27). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191-2194. <https://doi.org/10.1001/jama.2013.281053>
- Yigiter, A. B., & Kavak, Z. N. (2006). Normal standards of fetal behavior assessed by four-dimensional sonography. *The Journal of Maternal-Fetal & Neonatal Medicine*, 19(11), 707-721. <https://doi.org/10.1080/14767050600924129>
- Zoia, S., Blason, L., D'Ottavio, G., Biancotto, M., Bulgheroni, M., & Castiello, U. (2013). The Development of Upper Limb Movements: From Fetal to Post-Natal Life. *PLoS ONE*, 8(12). <https://doi.org/10.1371/journal.pone.0080876>
- Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezzetta, E., Scabar, A., & Castiello, U. (2007). Evidence of early development of action planning in the human foetus: a kinematic study. *Experimental Brain Research*, 176(2), 217-226. <https://doi.org/10.1007/s00221-006-0607-3>