Shapes and Staffs

A qualitative exploration of the ontology of computer-generated graphic scores

A thesis by **Pieter Pierrot**

For the consideration of the title Master of Science in Media Technology

Leiden University

Supervisors:

Rob Saunders - Leiden University
Edwin van der Heide - Leiden University

April 2022

Abstract

This thesis takes first steps into exploring the ontology of computer-generated graphic scores. Generating such scores with a computational system naturally fits the trend that graphic scores sparked in the mid–20th century, in which composers partly relinquished compositional power to the musicians performing their work. A generative computational system relinquishes even more control, as it has no musical ideas to communicate. This results in what I call a semantic vacuum. In this thesis, it is investigated what effects this vacuum has on the interpretational practice of musicians playing computer-generated graphic scores. This is done by qualitatively assessing aspects of their practice during live interpretation sessions, in which individual musicians sonify four computer-generated scores. The assessment shows that the influence of the semantic vacuum on the music-making process is quite small. However, reflection on different aspects of the system implementation and the experimental design suggests that complex relationships between contributing parties in the musical communication process exist that are not properly understood. In addition to the empirical results, the adopted theoretical and practical methodologies are discussed, as to build foundations for how to understand and study computer-generated graphic scores, as well as to advance the field of computational creativity by injecting humanities-inspired methodologies.

Keywords: computer-generated graphic scores, computational creativity, musical semiology, generative systems, interdisciplinary research

Table of Contents

I. Foreword

Ambition is a useful driver for any project only if it can be controlled and its accompanying goals are feasible. Knowing this, choosing a topic for my masters thesis was not a trivial task. The Media Technology programme prepares its students in this respect to make an independent judgment, as no topics are strictly off-limits. When every topic could in principle be acceptable for a masters research, one has to be chosen that has a balance between ambition and feasibility. Computer-generated graphic scores is a novel area of investigation that has received virtually no previous attention. Exploring such an area to discover what interesting questions and problems may be lurking can be called ambitious. The feasibility, or scope perhaps, proved to be something that had to be negotiated throughout the entire process, as possible perspectives on the matter and different questions kept arising. Each of these perspectives had to be considered; what would including/excluding it mean for the research? Certainly no easy task, but one that was necessary to get something truly interesting out of this study.

Practical matters made progressing the project quite difficult. The COVID-19 outbreak resulted in an isolated working environment, in which critical peer support and other social contact was missing. While online communication matured explosively and functions fairly well, I only met my primary supervisor Rob Saunders in real life the week before my thesis presentation. This lack of real-world feedback and communication is unnatural to academic practice in my experience. Then more than one serious piece of personal circumstances made it even harder to progress.

Nevertheless, after one year, here we find the final product. I am incredibly proud of the result. I learned while working on this thesis how to manage large-scale academic products and how to keep myself sane in the process. For me, the thesis proves that even in the craziest of times I am capable of soldiering on and keeping my standards high.

Without the nurturing and intellectual support of my friends and family this project could not have succeeded. In this respect I would like to thank my mother for helping me with both these types of support. During our walks I have told you more about the depths of this topic than I told anyone else, and your unconditional companionship cannot be overstated. Julia, your beautiful responses to the practical day-to-day aspects of my soldiering on never fail to keep my morale high and my feet grounded. A big thanks to Jesse, who helped me with a vital piece of practical feedback on the experimental setup. Bas, who always believes in me twice as much as I believe in myself and inspires me to explore. Then I would like to thank my musical collaborators, Baue Kunstman, Berend Eijkhout en Pieter de Mast, for blowing me away with your playing during the live sessions. Thanks to the Media Technology staff: Max, Maarten, Peter, Barbara and all the others, you all greatly helped me navigate my time during the programme and develop myself as an academic; keep producing high quality education and warm feelings. Lastly, I would like to thank my supervisors, Edwin and Rob, whose council was of essential value. Your seemingly unlimited time and effort kept me focused on producing academic work of high quality, while giving me the well-appreciated space to figure most of it out for myself.

Leiden, April 25, 2022

II. Introduction

Graphic scores are something of a hidden gem.

Most people, including but not limited to people with musical training, would recognize a piece of traditional sheet music immediately, with its characteristic five horizontal lines, the nice curly shape on the left of every bar and the blobs of black with the sticks coming out of them. This traditional notation system, culturally evolved over some hundred if not a thousand years (Treitler, 1982), depending on your definitions, is the dominant notational system in western composed music and is taught in and out of conservatories around the world. 'Being able to read notes' is an understood faculty for musicians: you don't need it, and excellent music can be made without it, but if you possess the skill, it opens up ways of representing musical ideas and communicating them to other musicians. That such a powerful and simple phrase, 'to read notes', has become synonymous with reading traditional sheet music illustrates the pervasiveness of this dominant notation system in musical communication circles. Indeed, far fewer people know that a tradition exists in which music is written in all kinds of other forms. These other forms can collectively be called graphic scores.

The apparent simplicity of the phrase 'graphic scores' is misleading: it has come to mean something like 'non-traditional scores', which is not a very informative phrase. If we stop for a moment to appreciate traditional notation as one option amongst many for communicating music, and look beyond its pervasiveness, we can see that a myriad of other options remain. The most obvious alternative is the use of some other notation system, with different visual characteristics and different rules and mappings from visuals to sound. But everything and anything can be used as a music score: photographs, flocks of birds in the sky, mathematical drawings, colored abstract shapes, world maps and stellar constellations; all these things and more have been used as musical scores[1].

In the first half of the 20th century composers began breaking down the traditional notation system. They felt the system had become restricting, after it had been providing them with a powerful compositional tool for the last few hundred years. This started with minor rebellions, such as adding previously unexisting symbols to their vocabularies or using symbols slightly outside of their traditional contexts (Cole, 1974). Reaching the 1950s and 1960s, however, with the notational and musical experiments of John Cage, Karlheinz Stockhausen, La Monte Young, Steve Reich and many others, graphic scores and new notational methods had become common practice. Such composers used all the things I suggested above and more to notate their musical ideas and changed the definitions of music in the process. One particular example: the direct inspiration for this thesis was a piece by Young which consists of one single horizontal black line on a white page (Young, 1960(2))[2]. Personally I think that such transformational ideas are artistic practice *par excellence*, and their importance should not be understated.

Academically speaking, graphic scores are a hidden gem, too. Since their integration into music practice in the 1950s they have persisted in a world where traditional sheet music continued to be the norm. During the last 70 years, composers have not stopped expanding the space of possibilities for communicating their musical ideas. At the same time, research on this topic is surprisingly sparse. Research that has been done is scattered across quite distinct areas of focus and perspectives, and is

[1] An informative collection of graphic scores can be found at the link below. This database has served as a general source of inspiration throughout this research.
https://llllllll.co/t/experimental-music-notation-resources/149

[2] Referenced scores can be found in Appendix B.

often accompanied by comments on the difficulty of researching graphics scores, because of their eclectic appearances and multifaceted nature.

This thesis will add a new line of inquiry to this scattered academic field. It will do so by taking some first steps into investigating the ontology of computer-generated graphic scores. In what follows, I will describe the process of theorizing, developing and putting to use a computational system that generates graphic scores. This system's intended purpose is to be a substitute for the traditional composer-figure: the human agent writing and notating the music. In replacing the composer with a computational system, I hypothesize that a crucial step in the musical communication process (from composer through performer to audience) will be compromised: a computational system does not have musical intention in the way a composer has, and this results in a lack of musical meaning in the resulting score. I call this lack of musical meaning the *semantic vacuum*. In this thesis I will explore whether computer-generated graphic scores will create such a vacuum and what effects it has on the meaning-making interpretative process that results in a piece of music.

The presented study can be seen as being constituted of two parts, although they are not presented in the text as being separate: a theoretical/methodological and a practical/empirical one. For the first part we will explore a number of theoretical perspectives to build some understanding of our barely understood subject matter. Because a lens through which to look at computer-generated graphic scores has yet to be fashioned, theory will provide perspectives on which to base our methodology. In turn, reflection on the adopted methodologies contributes greatly to theoretical discourse on graphic scores. The second part encompasses the system implementation and an empirical study, in which the generated scores are presented to musicians to be played. Both of these parts were progressed simultaneously as they both inform each other; we will see throughout this thesis how it is in this cross-influence that the most interesting discussion arises.

The outline of the rest of the thesis is as follows: First, in section III, I will provide some theoretical context on graphic scores. In section IIIa I will shortly describe the (art)historical developments that surrounded the work of Cage and his contemporaries and their intentions in adopting these new approaches to music making. In sections IIIb and c I will look at what research has been done on graphic scores and how we can build on that in spite of its scattered nature. These latter explorations will be grounded in work on musical semiotics and computational creativity, respectively. In section IV I will describe in more detail what is the area of focus of this study and with what methods will we investigate this area. In addition I will provide some clarifications on concepts that are used in this study. I will describe the system development part in section V, in which we will see how the theoretical framework that was constructed will help with conceptualizing the generative system. In section VI I will describe the experimental part, which consists of the live interpretation sessions. Being the crucial empirical part of the study, the sessions will provide a plethora of qualitative material to reflect upon. In section VII I will present results that will shed some light on the effects of the semantic vacuum. Then, in section VIII we will see how my conceptualization of the semantic vacuum holds up against the interactive setup that the experiments employed. I will discuss a series of reflections on the academic potential of researching computer-generated graphic scores, as well as a discussion on the used research methods.

III. Related Work

Before we can form a perspective through which to study computer-generated graphic scores, we will need some theoretical context. As any pre-existing framework for this area of inquiry was missing, this context was built from discourse in different disciplines. First, a description of graphic scores as a historical phenomenon will be given, that will provide some basic understanding of the things we are going to be generating with our system. Then I will get into some aspects of meaning and communication in music through work in musical semiotics. Lastly I will describe some work in computational creativity that can be helpful to our purpose. We will see, however, that still a lot of questions remain to be answered before we can get to the system implementation.

A. Historical context

The advent of graphic scores is as interesting as the phenomenon itself would lead one to suspect. It needs to be understood most importantly as a reaction to the directly preceding era of modernism and the position of the composer in that time. Before we can understand it as such, though, we will first need to take a step back to see how traditional music notation got its hegemonial position as communicator of music.

While the history of music notation is long and its key moments dispersed, a clear trend can be seen. Music scholars write about how the invention of musical notation and its integration in music practice shifted the focus away from the oral tradition of preserving and practicing music slowly (over the course of centuries) but surely to a written tradition (Treitler (1982); Magnussen (2019)). Before music was systematically written down, musicians always played[3] by heart and from memory; music was learnt through hearing and repeating. Such music was mostly monophonic and basic, consisting of simple melodies, so this did not pose serious problems (complementally, such ancient musicians often had insanely big catalogs of music remembered). As a result, these melodies did not need to be written down. As is often the case with oral tradition, the mouth-to-mouth (or instrument-to-instrument) way of communication meant that pieces of music were not stable, but rather evolving entities that changed over time with performances and as taste and practice changed. In this sense, one could barely write down such pieces anyway.

In an intertwining motion, music notation made music more complex, which in turn increased the demands for a powerful and stable notation system, which made writing more complex music possible, etc. With this development, around 1800 the concept of a composer[4] as someone writing the music as something to remain fixed took shape. The more intricate the written music became, the more power the composer gained over their work and its performance characteristics. What the composer notated constituted the piece and was to be respected. This resulted in what Umberto Eco (Eco, 1989) would call a closed musical work, which is made by a writing agent (classical composer or otherwise) and is finished and stable. Such stable works were already very much the norm in the times of, say, J.S. Bach, who wrote large intricate pieces that were carefully constructed and orchestrated and were (and are) to be played in a rather strict and normative manner. In the classical tradition, the composer's hegemonial power over their work has remained and strengthened during

[3] Playing should be read to include singing, which was by far the most prominent way of music making in these times.

[4] Along with other specialistic roles such as the conductor, the soloist, the orchestra player, etc.

years since then, resulting in even more complex normative music that was written until the first half of the 20th century. By then, composers reached a limit in both what audiences would accept while listening to music and what their vital instrumentalists were physically able to play.

The modernist culmination of this trend of increasing normativity and complexity is embodied by late 19th and 20th century composers like Gustav Mahler, Igor Stravinsky[5] and Arnold Schoenberg. Their music stretches the limits of the aesthetic, sociological and practical dimensions that constituted the classical music culture of their time. Inevitably, the musical objectivity sought after by such modernist composers proved to be impossible to attain; the human performers would have to be executing their musical (notated) ideas with superhuman accuracy. Philip Galanter writes (2009, p9):

> In modernity the focus of attention is on the author who, at the highest levels of achievement, engages in a high-stakes battle to create a totalizing theory and masterwork. [...] And when modernity reigned supreme over the arts, artists were similarly viewed as being singular, potentially heroic, and relatively unconcerned with their audience.

Now, in the 1940s such superhuman accuracy had become available with the advent of new technologies. Synthesizing sound had now become reality, as had recording. Musical objectivity had come to mean something entirely different. In addition to the human performer becoming a terrible tool for objective reproduction when compared to its own new inventions, with these technologies composers had new and exciting ways of making their music sound precisely like they wanted. Avenues of electronic music and other technological approaches to music making were explored by artists like Steve Reich, Philip Glass and Kraftwerk and have their own strands of influence of subsequent music, but our interest lies in what classical composers did to react to these developments.

In terms of actual music composition, postmodern composers had to think of something other than objectivity or complexity. After all, what was left to explore in terms of harmony, rhythm and compositional articulation after the music of their predecessors? While late modern and early postmodern composers (e.g. Britten, Bartók, Ligeti) experimented with extending their cherished notational system with other symbols and visuals that conveyed their musical ideas better (Cole, 1974, p132), in the 1950s and 1960s a new compositional strategy emerged that focused not on better communication tools, but communicating different musical ideas. John Cage, Morten Feldman, Earle Brown and Christian Wolff arguably being the very first composers of this new paradigm (Brown, 2009), but quickly followed by Karlheinz Stockhausen, Cornelius Cardew, La Monte Young and others, postmodern composers embraced the human imperfection in their performing musicians, and delegated parts of their compositional power to them. They started making music that was indeterminate, contingent, and relying on improvisation and interpretation. Earl Brown wrote about this (Brown, p6):

> I am not so much interested in the piece ultimately being a monument as I am in the piece existing as a kind of field of the activity of music-making which can exist between sympathetic and reasonable kinds of people.

A few examples can make clear what kinds of compositions might arise. Cage's *Fontana Mix* (1958) consists of some transparent sheets with lines and dots on them, which are supposed to be

[5] The premiere of his *Le Sacre du Printemps* in 1913, which instigated a small riot in the audience for being offensively modern, is a famous example.

overlaid in a way the musician seemed fitting, after which the emerging interplay between the shapes should be interpreted as a musical score. The musician has complete power over how to overlay the sheets and what the score ultimately looks like. Cardew's *Treatise* (1963-1967) is a 193-page book of abstract (although with hints of recognizable, traditional music symbols) lines, symbols and shapes, for the musician to interpret and perform. Earl Browne's *December 1952* (1952) is a highly abstract black and white score, consisting of black lines of varying length and thickness that are spaced out over a single white page. Young's *Compositions 1960* (1960), a series of conceptual text-based works, comprises a set of absurd instructions for the musician to 'play'. *Piano Piece for David Tudor #1* from this series reads, in its entirety:

> Bring a bale of hay and a bucket of water onto the stage for the piano to eat and drink. The performer may then feed the piano or leave it to eat by itself. If the former, the piece is over after the piano has been fed. If the latter, it is over after the piano eats or decides not to.

As noted, one result of this new compositional paradigm was that the power of the composer as the master of their work was partly being delegated back to the musician/performer. For Cage and his contemporaries, interesting ways of making music were to be found in factors that were outside of their control: (interaction with) the humans that actually sonified the music, performance-specific contingencies and actual chance. Their intent was not for their performers to replicate their scores as closely as possible to the best of their abilities, but to put in their own ideas, associations, beliefs about what music could and should sound like and how it should be played.

It is interesting to see how modern computers could be argued to have a special place on the power spectrum of a musical work. In previous paragraphs I have referred to new technologies meaning mostly early electronic equipment and computationally-oriented methods to composing such as serialism. We have seen how the development of a notation system expanded the complexities of the composer-role, and how this role was revisioned during the postmodern revolution under the influence of new technologies. Computers can be seen as a clear extension of this technological trend; consider their use in (generative) art and the academic areas of AI and computational creativity that are working on maturing the autonomies of computational systems. If we consider the computer as an enhanced version of the objectivist technologies of modernism and use it to replace the composer altogether, and relinquish the power of the composer to the musician-performer altogether, we can see how the semantic vacuum presents itself.

B. Theoretical context - Semiotics and humanities

In spite of the fact that graphic scores have been around for 70 years, parts of their ontology remain unstudied. That is, most research approaches the matter from musicological or (art-)historical perspectives. The contextual part above is written precisely from such a perspective. Scholars have put the mid 20th century developments into their historical and musicological context, to understand trends and the breaking with them. To put it bluntly: this they have done and are doing very well, and present thesis is not an attempt to complement this discourse. What we are interested in is aspects of meaning, communication and interpretation, as these are the areas of interest that relate closely to the semantic vacuum.

Graphic scores traditionally exist in between a composer and a performer, and function as conveyors of meaning (Casey, 2015). The inherent absence of an existing semiotic framework for these scores makes this process of communicating meaning far from straightforward. This is true for the standard composer-graphic score-performer relationships, but doubly so for our situation in which

the composer is replaced by a computational system. Therefore, if we are to learn about the ontology of computer-generated graphic scores, we will first need to have some understanding of the semiotic relationships that could be at play, and the different forms of graphic scores that communicate in different ways. On such topics, research is largely scattered. Given that our area of interest is in a niche within a niche, relevant writing is even sparser. We will look at a few approaches that can be of value to our purpose.

In musical semiotics, some theoretical and experimental work on graphic scores has been done. Semiotics, the study of signs and their referents, informs us about relationships between score and music, and the way meaning is transferred between the composer, the score and the musician. The theory of Peircian semiotics has been applied to sounding music (Tarasti, 2002) and musical scores (Monelle, 1991) to tackle questions of relationships of music and the real world and the way music is understood. Inspired by such work, George Athanasopoulos asked if the success of applying such semiotics to music could indicate that a shared semiotic background between individuals (e.g. western culture) leads to similar interpretations of an alien notational system or a graphic score (Athanasopoulos, 2010). He does this by presenting abstract scores of his own making to musical performance students and qualitatively studying similarities between their performances. Indeed, he finds that such similarities exist. Virginia Anderson (Anderson, 2013) describes such similarities through her concept of *sonic identity*, the ontological and perceivable intersection on which all performances of a particular work exist. What this tells us is that, even in abstract scores, a shared sense of musical identity of a work (a score) can emerge without complex, semantically rich notational structures. While an identity of a musical piece is not precisely what we're concerned with, the notion does naturally extend to the fact that music can and will be read from abstract scores that are freely interpretable. This is hardly a surprise, considering the impact of Cage and others on the history of music, but it justifies our search for insight into situations in which the initial meaning-maker, the composer, is stripped away: it challenges our assumption that the absence of an intentional composer necessarily bars the potential for musical meaning.

How similarities between performances of abstract graphic scores hint at potential semiotic meaning in these scores is interesting, but it concerns more specifically the question of what constitutes a musical work or an instantiation of that work, a question that only tangentially helps us. Rob Casey uses the theory of semiotics to argue for a phenomenological perspective on the way graphic scores communicate intentions. For Casey, interpreting a score is an embodied process, in which all bodily and worldly factors play a role. The difference between traditional notation and graphic notation then is not so categorical, as for the latter performers will use layers of dynamic understanding (that closely relate to their understanding of traditional notation) to interpret the abstract graphics (Casey, p170):

> "[the score] is not explicitly a signifier of concepts, actions or objects but a purveyor of affect. Interpreting a graphic score by translating the image into metric data and assigning musical properties to that data reaffirms the assumption that the sensory information issuing from a static image conveys nothing without the viewer's abstract reasoning to complement it."

Here, Casey implies that graphic scores need no underlying existing framework to be understood; understanding will be imposed by our biological circumstances. It is our embodied perception, our proprioceptive understanding of our worldly bodies, with the score, the room, the instrument, the audience, that informs the performer's perception of the score and their meaning-making process.

Both Casey and Athanasopoulos, then, would be skeptical about the existence of the semantic vacuum.

Such lines of reasoning lead directly into the work of Lambros Malafouris, who argues that when it comes to embodied interactions with real-world objects, those objects influence us just as much as we influence them (Malafouris, 2008). Malafouris calls this 'material agency': "If human agency *is* then material agency *is*, there is no way that human and material agency can be disentangled" (italics from original). His original example of the potter and the clay is clear enough: without the clay being there pushing back on the hands of the potter, no pot can be made. Graphic scores have such material agency in the sense that they shape the performance, by existing in a constant interactive feedback loop with the performer. They exist within a coupled system that comprises the score *and* the human. Such theory suggests that a computer-generated graphic score can convey meaning 'just' by entering a relationship with a human performer.

C. Theoretical context - Computational creativity and analytics

Theoretic work on musical semiotics and meaning is helpful and comprehensive, but does not bring us very far in constructing a framework for *computer-generated* graphic scores. To build such a framework, we will have to leave the qualitative musings of the humanities and take a look at the fields of computational creativity (CC) and new-media interfaces, in search of more analytical work and computational systems that could inform our strategies to move forward.

Research on actual computer-generated graphic scores is virtually non-existent. Then, most CC research focuses on the assessment and judgement of computer-generated creativity (Colton & Wiggins, 2012). Such work is generally not very relevant for our discussion, as the artifacts produced by our system will not be judged as creative artifacts. In addition, some work has been done on the development and analysis of interactive tools for graphic score creation (e.g. Adhitya & Kuuskankare (2012); Mattinson & Sarkar (2020)). Such systems are to be used collaboratively by human users to enhance and complement their own creativity. While on the surface this looks useful, 'users' of our system, if our performing musicians could be called that, can not interact with it. While its artifacts are to be considered as having material agency and are to be related to in a bi-directional way, the system remains autonomous in this sense and does nothing to accommodate this relation. Also, interactive systems are often developed from an interaction design perspective, which focuses on user evaluations and iterative design. Because we opt to attempt a clean swap between a composer-figure and a computational system, such an approach will not work.

One piece of helpful analytic work has been done by Brian Inglis (2015). He recognizes the lack of systematic research on graphic scores, writing how they "[provide] continuing relevance for composers and performers and a largely untapped phenomenon for music analysis" (Inglis, p11). He builds upon the writings of Virginia Anderson to construct a typology of graphic scores, which can be found in figure 1. This typology is a humble but thoughtful attempt at organizing what is otherwise an eclectic mess. Inglis does not suggest, and nor do I, that this model is perfect; edge-cases or category-defying scores are not hard to find (consider, for example, Young's *Compositions 1960*). It

I Graphics lacking musical signs	II Graphics incorporating some musical signs	III (Determinate) graphic elements hybridised with conventional signs/meaning
1 single page		
2 multiple pages		
a) No axes ('pitch and duration lattice') implied	a) Not (necessarily) usual symbolic meanings	
b) Axes implied	b) Usual symbolic meanings	
i) No verbal information (instrumentation/ instructions/explanations)		
ii) With verbal information (instrumentation/ instructions/explanations)		

figure 1 - Graphic scores typology

could of course be argued to be senseless to try to taxonomize a phenomenon as fluid as graphic scores, which likely is the reason so few attempts have been made. In most cases, however, the model does its job and as we will see in section V, this piece of systematicity is precisely what we need when we navigate the problem of what the system and the scores should be like.

In the same paper, Inglis extends Nattiez' and Anderson's semiotic model of the composer's communication of their intentions. Inglis argues that there are two poietic processes going on in transmitting a musical idea: from the composer, through the score, to the performer and from the performer, through the performance (or it might be some other channel, such as a recording), to the listener. The model that Anderson and Inglis propose place the performer in an 'esthesic' as well as an 'poietic' position, meaning they semiotically 'take in' as well as 'send out'. The taking in is happening when they read the score, basically, or otherwise learn about the intention of the composer. After this stage, they form their own intentional transmission of ideas, namely in the way they perform the piece. What happens between (and including) the esthesis and poiesis of the performer is what is under scrutiny in this thesis.

An area that has attracted recent attention in CC communities is that of human agency in CC systems design and the effects this has on the alleged autonomy of computational systems. Bidgoli, Kang and Llach (2019) argue for the importance of recognizing the web of (human) interactions that surrounds AI-enabled tools. As the authors press, such tools act as 'machinic surrogates' for their human creators, representing their proposed form of interaction with all the details of implementation that come with it. Interestingly they apply their theory to actual co-creative tools in the context of art and creative ownership; these systems are generally designed to be interacted and co-created with, which changes the relationships at play when compared to our composer-performer relationship. The perspective these authors take on the design of computational systems and what their perspective could mean for our generated scores will prove to be an important point of reflection; we will see why in the system implementation section and we will return to the implications of this perspective in the discussion section.

IV. The Study

We have seen how some pieces of theoretic work can be helpful in understanding our area of inquiry and constructing a framework through which to implement the generative system. In this section I will describe in more detail what this study is about. The first part describes the adopted methods; the second part concerns the concepts that are important to clarify.

A. Methods

The semantic vacuum is a hypothesis of sorts. It is not necessarily grounded in theory, but rather the result of the basic understanding of computational systems as lacking consciousness or intentions; a basic understanding that served as a starting point for this research. Such a starting point was necessary to have, because it was the only way to enter into unexplored territory in which every perspective is a valid and possibly fruitful one. It is because of the evident nature of this understanding that the semantic vacuum is useful as a lens through which we can look at the ontology of computer-generated graphic scores. In addition, a quick and dirty comparison between a composer of graphic scores and an imaginary score-generating system implies a vacuum: 'real' graphic scores themselves can be easy to generate, as the La Monte Young score from the introduction illustrates. The interesting part is that generated scores are never *intended* to be musical scores in the way that any graphic score made by a composer is. As such, the vacuum reveals itself as a missing link in the musical communication process. The phrase "vacuum" has a useful metaphoric quality in this sense: if it would exist, what would happen to its surroundings (the performer, the music); will they implode and crash, or naturally and actively fill up the space that the vacuum leaves?

To investigate the effects of the semantic vacuum, a computational system was built that generates graphic scores. The scores were then presented to three (semi)-professional musicians, who interpreted and sonified the scores. This way, the scores were investigated as potential vehicles of musical meaning that is transmitted between the composer and the performer. In the live sessions, I qualitatively observed the musicians as they interpreted the scores, and conducted a semi-structured interview that inquired about their experience working with the score and the musical choices they made in the process. Reflection on these observations will inform us about whether a semantic vacuum exists, how it exists and what could be implications for new ways of music making and doing research.

In addition to its grounding in computational creativity theory and practice, the study employs a curated form of ethnographic research, which can be understood in two respects. Because the area of computer-generated graphic scores has not yet been explored, a straightforward approach to studying it was not available at the start of the research process. The objective of studying computer-generated graphic scores was set, but what aspects of the topic to study and how to study them were not. Instead, a method had to be devised that allowed for freedom to explore different perspectives on the topic and to reflect on what these perspectives could contribute to our understanding; an ethnographic approach works for this purpose. Effectively, this approach consists of respecting that different strands of research practice need to be progressed simultaneously and recognizing the influences between these strands. The different parts of practice are: theoretical study on the phenomenon of graphic scores and the topics of meaning and communication and music, the development of the computational system and the construction of an experimental setup that would fit the research objective of learning about the ontology of computer-generated graphic scores through the semantic vacuum.

Traces of this ethnographic method can also be seen in the thoroughly qualitative data gathering approach. As we are interested in learning about the effects of the alleged semantic vacuum, a novel perspective on musical communication and meaning, we need to explicitly leave some room for findings that were not anticipated. No preconceptions of the vacuum can help us with constructing a rigorous and well-defined boundary of the topics to study. Instead, we will see how the live sessions setup and the semi-structured interview provided the musicians with room to present their personal experiences of interpreting the scores, with the intention of following a natural assessment process that allows for serendipitous findings.

As the construction of a suitable methodology was of great importance to the success of this research, reflection on the methods and the choices that were made to make progress towards the research objective will prove to be a relevant contribution to discourse on graphic scores, as well as to discourse on perspectives on computational creativity. We will see that pivotal choices were made in the system implementation and the experimental design to steer the study in the most fruitful direction. In the discussion section we will reflect on these choices, and what other potential perspectives exist.

B. Concepts

Here, I will expand upon some central concepts of this study. This entails conceptual clarifications as well as descriptions of concepts that were formed and developed as vital parts of the research. To this end, it is important to note that while the semantic vacuum has, in some form, always been the central area of interest, it has not always meant one static thing. It is better thought of as some distant silhouette that changes shape and gains detail when approached. The concepts described in this section, then, are best understood as some necessary stepping stones to get that silhouette into focus. That they are presented as stepping stones does not mean they were formed in this order chronologically. Reconceptualization was needed throughout the research phase and up to the live sessions; it is exactly for this reason the curated ethnographic approach was adopted.

1. Musical meaning

Meaning in music is most familiar to anyone who enjoys listening to music, or making it. It is that quality that makes it worth listening to, the part that speaks to us. Much can be said about the ways music can speak to us, from a cognitive level, a semantic level, a physical level, and so on. Importantly, however, is the fact that meaning in music is primarily experienced through its manifestation as sound. As such, it is quite different from the way meaning is present in musical scores. In a most banale way René Magritte would have been proud of, a musical score is in itself just a piece of paper. Its meaning is not in the symbols or other visual elements that are written, just as the meaning of written text is not within the textual symbols themselves. Written musical elements have only a faint resemblance (e.g. Inglis' time and pitch axes) to the music they code for, after all. No music lover would get emotional after seeing a beautifully and masterfully crafted piece of sheet music.

Meaning in musical scores is not the same as meaning in music: scores get their meaning from their relationships with the composer and the musician. More accurately: scores are *vehicles* of meaning that transmit between the composer and the musician. The composer transmits their musical ideas through presenting them in a visual format, be these ideas about emotion or about redefining musical relationships, and the performer interprets this visual material, be it traditionally notated or notated graphically. Of course the construction of musical meaning on both ends of this transmission can vary wildly depending on the form of the ideas and the vehicle. A fugue by J.S. Bach will be

understood quite differently than an aleatoric score made by Cage, for example. In either case, however, the score functions as a transmission medium between the composer and the musician.

Then, what I mean when I use the word 'meaning', is the transmission of intentional information. Meaning is not in the score itself, but in the mutual understanding between the performer and the composer that the artifact is, indeed, a musical score, and codes in some way for music to be made. In most cases, the idea for a new piece of music comes from the composer, who intentionally writes down their idea to communicate it to someone who can sonify it. The intentional information flows from the composer to the performer, the latter then starting their own transmission of intention through the music to the audience.

An important distinction has to be made between how this musical idea is represented and can be extracted from graphic scores on the one hand and traditionally notated scores on the other. The main difference, of course, being that the latter is embedded within a rich cultural tradition that creates common ground between two users of the notational system. 'To read notes' means the same thing to those users, because elements and structures in the scores have the same rules and semiotic referents. Composers and players of graphic scores do not necessarily have such common ground, which is inherent to the experimental and experiential nature of such scores. That is, common ground will not be found in the visual aspects of the score. It might be found in any implicit or explicit rules between composer and performer concerning what is allowed or encouraged during the interpretation process and performance. Such rules or instructions are often provided by a composer to create some space of performance possibilities, which I would call common ground. However, such instructions rarely carry the amount of specificity needed for a straightforward translation, as a focus on interpretation is central to graphic score practice.

2. Score-performer relationship

This thesis focuses solely on the relationships between the composer, the score and the performer, and the meaning-making process that takes place between them. It is believed that the most insightful process for investigating the ontology of computer-generated graphic scores takes place within this triangle. It does for the most part leave out the important relationship between the performer and the audience, and with that the actual music that is made between the two. From the perspective of computational creativity, including the latter relationship would have made sense. Lots of work has been done on systems that create music and debate about the assessment of such music is plentiful (see, for instance, Carnovalini & Rodà (2020)). Entering such debate would have made it easier to ground the study in other work, for instance by providing a way of assessing the generated artifacts. Furthermore, it could be argued that researching musical scores without focusing on the music that is made from them makes no sense. After all, musically speaking, the interesting part is not happening between the composer and the performer; if the process stopped there there would be no concerts or recorded music, no actual music to listen to.

However, choosing to focus on the composer-performer relationship does not inherently mean that the sounding music is redundant. It only means we will be leaving out the *audience*. The score is of course inextricably linked to the music it codes for, and it is vital that the music is actually being made by the musicians. They couldn't be asked to explain their interpretation strategies without first being given the opportunity to form such a strategy, or something of the sorts. One can investigate the score through the music without letting the music exist for an audience per se. From this perspective,

it makes more sense; the audience, after all, has rarely anything to do with the musical score[6]. Keeping in mind the boldness of the choice of leaving the sounding music out of a study into musical scores, we will return to this topic in the discussion section.

3. Affordances

Before the system implementation started, the semantic vacuum was mostly seen as a problem for what the generators should be and what the scores should look like; if there is no composer who decides what to notate, then who decides? Certainly the answer to this question should not be a simple 'me, the system designer', for this would clearly render the claim of the scores being 'computer-generated' less than legitimate. It would mean that I would be making the scores *on* a computer, rather than let them be made *by* a computer, and hereby betray the spirit of the study. The vacuum was seen as something to be overcome: it needed to be filled with some semantic content before progress in understanding could be made, of which the author could not be me.

One of the most important tools to fill the vacuum was conceptualized as the scores' affordances, after a theory by William Gibson (Gibson, 2014). Originally devised to be used for describing action-possibilities in natural environments of animals, the term has been injected into a variety of fields, among which human-computer interaction and computational creativity. In such contexts the term has come to mean the action-evoking characteristics of machines or devices. A button affords pressing, a folder affords opening by clicking, etc.

For our purposes, the affordances refer to the characteristics of the score that evoke certain musical reactions in the performers. They are the visual elements the score is made of, and which the musician can latch on to. Asking about a score 'what does it afford?' can be a field of research in itself, as the eclectic and experimental nature of graphic scores should make clear. In this study, I will adopt a simple heuristic that comes down to basic gestalt principles, such as 'less vs more elements' (e.g. Young's horizontal line score has little affordances, a score like Cardew's Treatise affords a lot more), as well as some more semantically driven principles (e.g. abstract shapes vs notes), that I will come back to in section V.

It was recognized early that the scores' affordances could provide some useful guidelines for the system development and generation parts of the study. The problem of what to notate was partly solved by focusing on the visual appearance of the score: if meaning does not come from the (intentionless) system and cannot be added by me, maybe meaningful interpretation can be provoked through the affordances of the scores. This could be called presenting a suggestion of meaning, in the absence of actual meaning.

4. Context over content

The phrase 'context over content' could be seen as a higher level generalization, or integrated whole, of some separate sub-conclusions concerning meaning-making and the musical interpretation process. Specifically, it is the result of the focus on the score-performer relationship. In this phrase, context refers to the way the scores are framed and presented to the performer, while content refers to the visual aspects of the score. The latter contains the actual shapes and staffs, the printed material, but also the more structural aspects such as categories of vocabulary and syntactic structures. Context *over*

[6] This is less true for graphic scores than it is for traditional scores. Lots of graphically notated pieces blend the roles of performer and audience, something that can be said of postmodern art in general. A modern instance of such pieces is Metaphysics of Notation by Mark Applebaum (2008) (Applebaum, 2008).

content, then, means that for this research the framing and other more general circumstances that surround the score are more important than what is on the page.

This assertion is grounded in the struggle between the urge to fill the semantic vacuum with complex visual material from advanced AI-techniques and the suspicion that trying to force some semantic framework into the generation of the scores might compromise their computer-generatedness. During the first stages of the research, I envisioned the scores as being created by some Machine Learning (ML) technique, as systems that employ such techniques can create novel and interesting visual artifacts. A ML approach would have come down to collecting a set of graphic scores and using them as input for my system, which would then generate an artifact based on this input. More generally, computational creativity research conceptualizes the collection of material that served as inspiration for the researcher, which does not even have to serve as actual input. This is called the Inspiring Set (IS) (Ritchie, 2007, p76):

> "The construction of the program is influenced (either explicitly or implicitly) by some subset of the available basic items. This subset, which we will call the inspiring set, could be all the relevant artifacts known to the program designer, or items which the program is designed to replicate, or a knowledge base of known examples which drives the computation within the program."

Now, finding a suitable IS for graphic scores (note the lack of specification) comes with some problems. For one thing, any handpicked IS will have a large impact on the possible generation space that comes directly from my human influence, making the computer-generated nature of the scores questionable. In addition, picking any subset of scores and using it as models for our generated scores could result in artifacts that either do share some perceivable visual framework with a score from the IS, or don't share one. Both of these options lead into problems with deciding how to generate the scores. The former would betray our investigation of the semantic vacuum, as a compositional approach would be copied, to some significant extent, from scores in the IS. The latter would lead to arbitrary artifacts that bear little resemblance to any existing musical score. In this case some significant cherry picking would be needed, again increasing my human influence on the generation process.

In addition, I was skeptical about the use of increasingly complex implementation techniques in the generation process to work around the semantic vacuum. For taking first steps in exploring some topic, simplicity helps with focusing on the area of interest and eliminating the possibility for alternative explanations of the observed facts. Furthermore, elaborate computational processes, under which I would think ML techniques to be, cannot create semantics, only the illusion of it. Building a Generative Adversarial Network-based system, for example, to create visually striking scores that resemble the scores that served as input, would have been a cheap way to apply a semantic coat of paint to an otherwise equally meaningless artifact.

The 'context over content' paradigm approaches the semantic vacuum from the other way around. Instead of trying to fill the semantic vacuum ourselves, as part of the system implementation approach, the key is to let it exist and observe it as it exists in a musical interpretation context. This paradigm acknowledges the semantic vacuum and tries to study it in full daylight. In this sense, I abandoned the system-driven perspective that leads us into problems with what to generate, and turned to a more process-oriented perspective that focuses on the actual interpretation process. While the paradigm did not entirely solve the problem of what to generate, it made the problem a lot smaller and easier to navigate. After all, a focus on 'how' the scores are interpreted, instead of on the 'what' is

being interpreted in the score, provides simpler strategies for the implementation of the system that satisfies our needs. We will see in greater detail what approach was chosen in the next section.

A focus on the context of the scores was accomplished through the framing of the study to the musicians and the presentation of the scores in the live sessions. Through this framing a context for the scores was constructed that encapsulated the score, the system and me, the designer of the system. Indeed, this came down to simple phrasings and choice of words, as there is a significant difference between 'I designed a system that generates these scores' and 'These scores were generated by a system' for the performer's understanding of what the scores are and what parties might have been involved in their development. I will go into more detail about how the scores were framed in section VI.

V. System Implementation

Inglis' typology is helpful when we look at it from an implementation perspective. Even though we have up to this point only simplified the problem of what to generate by adopting the view that through precise framing we can avoid using advanced computational techniques, upon inspection Inglis' categories provide us with all we need: the categories themselves can serve as the 'what': scores can be generated that fall right in the center of a respective category and are designed to be just that. A closer look at the typology shows, however, that not all the categories are helpful: generating a 1-page vs a multi-page artifact is of arguably no relevance to our purpose. The section of the typology that I use is highlighted in figure 2. For convenience purposes, I renamed the categories into C1 through C4. This modified categorization can be found in figure 3.

I Graphics lacking musical signs	II Graphics incorporating some musical signs	III (Determinate) graphic elements hybridised with conventional signs/meaning
1 single page		
2 multiple pages		
a) No axes ('pitch and duration lattice') implied	a) Not (necessarily) usual symbolic meanings	
b) Axes implied	b) Usual symbolic meanings	
i) No verbal information (instrumentation/ instructions/explanations)		
ii) With verbal information (instrumentation/ instructions/explanations)		

figure 2 - Used categories from Inglis' typology

The categories of the framework that concern verbal information (the bottom two rows in figure 2) were deliberately left out of the study; primarily for simplification reasons. In artificial intelligence and computational creativity research, natural language is a field on its own. After all, language is human's most powerful tool of communication. Were we to add (generated) instructions to the scores, we would lose our grasp on what we are studying, as they would flood the musical communication process with ways of making meaning through relationships between text and picture. Moreover, adding instructions would add the difficulty of having to build a system with two major components, a textual component and a visual component, that have to be coherent enough to create an artifact that is suited to be studied in this context. As graphic scores often (but not always) do come with instructions, for instance a legend or a set of rules, leaving them out of this study contributed to the neutral framing of the generated scores that was employed. We will reflect on this deliberate choice in the discussion section.

The four categories that remain, then, together comprise a useful set of scores with different levels of semantic content. C1 through C4 can be thought of as having increasing levels of semantic complexity, the former being the most abstract and the latter being the most explicit and rich. C4 will have some resemblance to actual traditional music notation, while C1 will have only abstract shapes. Differences in these levels of semantics can be used to study differences in interpretation strategies, and inform us about the semantic vacuum and how it could manifest itself in relation to the musical meaning-making process. Although thinking of these levels of semantics as lying on a spectrum would be misleading, as the differences between the categories are, well, categorical, they can be thought of as having increasing semantic content.

C1	No musical signs, no axes
C2	No musical signs, with axes
C3	With musical signs, without their usual meaning
C4	With musical signs and their usual meaning

figure 3 - Renamed categories

In visual programming environment Processing, four separate generative programs were made. Each program generates, upon running, a unique score that is drawn on screen and saved as PDF. The final versions of these scores can be found in Appendix A. The programs were designed to produce scores that fall unambiguously within each of the four categories C1 through C4. This itself is quite straightforward, as the categories are distinct enough and their respective characteristics clear.

The necessity of producing scores that fall unambiguously in a specific category, however, did present a problem with the necessary variety in the generated scores. In other words, each time the program is run, a score should be generated that falls in its category, while being unique enough to still be called computer-generated. That is, I could have implemented programs that create scores to fall in a certain category without any autonomy of the system; those scores would be exactly the same every time. This would change the scores from being made *by* computer to being made *on* a computer. Generating within the space of possible scores (per generator) that satisfies both the usefulness of the categorical framework as well as the needed variety between individual scores was an act of balance that makes us recall the machinic surrogate discussion presented by Bidgoli et al.

The results of this can be seen in the elements that make a score fall into its respective category. The abstractness of C1 was realized by using only lines and circles, abstract shapes that can be interpreted in any number of ways. The x- and y-positions of each line and circle are used as parameters for shade, size and thickness, each of which in turn can act as an affording dimension of the score. In C2 the horizontal spacing of the score, as well as the 'H' and 'L' and numerical indicators, imply the pitch and time axes, while the shapes between the horizontal lines, as well as the additional lines underneath each system, are shaped and spaced randomly. In C3 we see the traditional staff and clef, which also imply pitch and time axes, as well an ongoing line on the staff with circular and rectangular events. We see a mixture of traditional notational material with other non-symbolic elements, as constitutive of the category. Importantly, while the traditional staff and clef are used, they are not used in their usual context where time would be measured by bars and pitch is scaled discretely and systematically. C4 spaces some staffs with clefs and notes on random positions, with lines connecting some of them. The notes on the staff are also randomized, which means the resulting harmonies are the result of chance and not composition.

As should become clear, the strategy I used most prominently to ensure generation variety is the use of randomness. The contents of each score are randomized to a degree that I, as programmer, could not have predicted the exact outcome. Unpredictability is a necessary but not sufficient prerequisite for a generated artifact: each artifact should have some degree of unpredictability to the programmer. This is needed to ensure the generatedness and rule out the possibility that the programmer is actually designing the thing behind the scenes and covering up their tracks, knowingly or not. About the part that plays outside of this unpredictability, that is, the part that *is* predictable to me as programmer and that could be argued to be the product of design, a few more things should be said.

The problem of assessing the autonomy of computational creativity systems is relatively new. Because of the necessary influence of their human designers, who have to make all sorts of choices during implementing a system and thereby defining its generation space, these systems are never fully autonomous. Naturally, discussions arise about creative agency and the lack thereof, and about the interplay between the systems' influence and their designers' influence on the artifacts. In this research, such discussions are highly relevant. As one of the tenets we're concerned with is acknowledging and centering the lack of the system's creative or artistic musical ideas, watchfulness of letting one's own creative or artistic musical ideas take center stage is crucial. This would betray the generated nature of the artifacts, and with that the central aim of this research. This is easier said

than done, of course, precisely because of the unavoidable part I as system designer play in the generation process.

In the implementation stage, the main strategy to navigate this problem was to never program something with a possible musical translation in mind. Elements that were programmed, such as the shapes between the horizontal lines in C2 or the lines connecting the staffs in C4, were not programmed for any musical purpose. Part of this mindset means relying on the workings of randomness, but for some elements this works better than for others. From the fact that the musical notes in C4 are randomized on their respective staffs, for example, it can be argued that it lacks harmonic intention. The gradient in C1, however, is programmed as a gradient, and will remain one in spite of the randomized elements that constitute it. In any case, even such elements were programmed without any ideas about how they should sound, or how they should be played, or on what instruments.

Then, there is an important link between the scores' appearances and their affordances, that played a role in the software development stage. That the framing of the scores is more important than the visual contents of the scores does not mean the contents are of *no* importance. What is produced, visually, is all musicians have to work with. While developing the systems, therefore, the focus was on scores that have a decent amount of affording elements, and that this amount is somewhat consistent over the different generators. In effect, we see in each score some affording elements. In C1, the lines and circles themselves, but also the differences between them: length, size, shade and so on. In C3, the curve as well as the thickness of the line, the three parallel lines with the shaded circles, the rectangular and circular shapes, and so on. All of these elements afford being objects for musical interpretation.

One question that I had to ask myself while implementing was: when is a score finished? Even though randomness and the absence of musical focus during the implementation brought us quite far in building the systems, a choice that had to be made unavoidably was when to stop changing the generators. I used a small set of stopping criteria when addressing this problem:

1. *The score is appropriate for its respective category*
   As noted, because of the clear distinctions between the categories this is quite straightforward.

2. *There is sufficient variation in the scores to afford creative interpretation*
   The heuristic for this criterion is the number of different affording elements that is in the score. For instance, in C1, we have: horizontal lines, vertical lines, circles on either end of the horizontal lines, the greyscale of the circles and the strokeweight of the lines. In C3, we have: the height of the main line, the thickness of that line, the greyscale of that line, the circular and rectangular events and the spots where the line splits into three and is accompanied by the tunnel-like visual.
   Each of these elements affords interpretative freedom and choices.

3. *The artifact feels algorithmic*
   This criterion has close ties to the general mindset of not putting my own creative ideas into the scores.

4. *The artifact looks finished*
   Assessing this criterion is largely based on a set of judgement calls: is there anything missing from the scores? Is the canvas space used properly, i.e. are the elements

spread out over the page? Is there any messy overlap between elements? Passing this criterion came down to some amount of cherry-picking the artifacts: because of the randomized nature of score production, some generations just look wrong.

Important to note is that none of these criteria concern the translation of score to music. Criterion 2 does concern the translation process, but it is merely about the perceived affordances of the score, and not about how the affording elements could sound or be played on any instruments. In addition, note how none of the criteria are concerned with whether or not the score looks *good*. My aesthetic judgment (or that of others, for that matter) was of no importance to calling the scores finished.

VI. Experimental Setup

In this section I will describe the setup of the live sessions and expand on the methods of documentation that were used. I also dedicate some space to the introduction of the collaborating musicians.

A. Live sessions

The live sessions consisted of an introductory interview, a performance part, a video playback part and an in-depth semi-structured interview. The purpose of the introductory interview was to get some sense of background on the musicians and whether they had any preconception of what the session was going to be like. Above summaries were constructed from these interviews. The introductory interview and the semi-structured interview can be found in Appendix C. The performance part was recorded on both video and audio. In addition, during the performance part, notes were taken that captured any general observations about the process, such as hesitation, musical flow, extensive looking at the score, silence, etc. These notes were used as guidelines in the video playback part. The interviews were recorded on audio, and additional notes were taken in real-time.

In the performance part, the musicians were first presented with a practice score to play, for which Earle Brown's *December 1952* was chosen. This highly abstract work was deemed appropriate for a warming-up exercise, to prepare the musicians for their active role as interpreter, as well as to provide an open environment for musical warm-up. Through the thoroughly abstract nature of this score, the interpretative power of the performer was intended to be drawn out and highlighted immediately. After the warming-up, the musicians were presented with the four computer-generated scores, one at a time. The order of presentation was randomized, as to eliminate any pre-conceptualized or structural cross-score influence. The musicians would perform the scores without any substantial breaks in between (i.e. no time for reflection, questions or comments).

After the performances, the musicians were asked to give an account of their initial thoughts and experience. Rather than to force the musicians into a rigid interview setting right after playing, it was deemed more natural and fitting to provide the opportunity for them to articulate their untainted experience. Some topics that were scheduled to be covered during the final interview, such as which scores were harder to play and the visual affordances of the scores, were already touched upon in this part.

The video playback sessions were implemented to provide a more reliable and substantial source of reflection from the musicians. During the performance, their attention had been on reading and playing. However, I wanted to get a personal account of what choices were made while playing, and how those choices translated into sonic material. A retrospective playback session was deemed appropriate for this purpose. During this playback, interruptions made by either of us were welcomed, to comment or ask about certain parts of performance, such as timbral choices, choices about translation or interpretation and quality of music.

The final semi-structured interview was designed to cover all the topics of interest. If some parts had already been covered in other parts of the session, they could easily be left out in the interview. The interview was structured to discuss topics of increasing relevance and depth in three categories:

- **Affordances**. What was there to play? Was it hard to interpret the scores? What did the musicians think about the different aspects of the scores and their affordances? Questions on

this topic are about the visual aspect of the scores and how music was made out of shapes and elements.

- **Computer-generatedness**. Was it noticeable to the musicians that the scores were generated? Did this influence their performance? How did they perceive the presence or absence of meaning in the scores? Through such questions, I intend to learn about the relationships between the system, the score and the performer, as well as musical meaning or the lack of it in the score.
- **Composer role**. Did the musicians have ideas about what computer-generated might mean in this context? Did they have ideas about who might be called the composer of the scores, or what might have happened to this role? These questions address the most high-level aspects of generative scores and their ontology as vehicles of meaning between the intentionless system and the performer.

Most conclusions that we will visit in the next section were drawn from information that was learned in this interview.

As described in section IVb, it is in the experimental design that the framing of the scores becomes of great importance: how the scores are presented to the musicians determines how they will relate to them. Because we are investigating the relationships that could be at play between the score and the musicians, I chose a very natural way of presentation, as to not tarnish these relationships any more than needed. In practice, this means that the musicians were told beforehand that they were going to interpret computer-generated graphic scores, in this exact phrasing, but not how these scores were made or what they looked like. The same message was sent to Berend, Pieter and Baue to prepare them for the sessions, which included the above statement as well as some practical information, such as that they were free to choose their own instrumentation.

An important aspect of the framing of the scores is the fact that the musicians did not get any preparation time in advance. That is, they got to see the scores for the first time during the session, right before they were asked to play it. This approach was chosen because of what we are focusing on to learn about the ontology of computer-generated graphic scores: the interpretation process of musicians. Were we to send them the scores beforehand, which arguably would fit into a more natural process of rehearsing and performing music, we would have no means to observe this process. By putting the moment of first contact and the performance this close together, I intended to evoke performances that are direct and experiential.

In fact, this part of framing the scores was inspired by a compositional approach that was used by many postmodern composers during the last decades. This approach occupies the most free and performer-oriented side of the graphic score compositional spectrum. In this paradigm, scores are often highly abstract and written without any specific instrumentation or instruction. An example of such a piece is Applebaum's *Metaphysics of Notation*. This piece is designed to be played by whoever encounters the score, without any preparation time and without any instructions as to how the piece should sound or be played. In effect, there are no ‘wrong interpretations’ as far as the composer is concerned (Applebaum, 2008). Keeping in mind the rather quick and experiential interpretations of the scores I intended to evoke during the sessions, eliminating the danger of the musicians ‘playing the piece wrong’ is a good way of framing the scores in a natural way, as well as providing a safe environment for the musicians to play. Nonetheless, we will return to this topic in the discussion section, to reflect on what this lack of preparation time means for our study and what could be learned from adding it to an experimental setup.

B. Collaborators

For the live sessions, three musicians were selected for their skill and creativity. Skill is necessary, because I wanted the musicians to have a wide range of techniques on their respective instruments, to be able to be free in their translation. The more skill, the more options and the closer the musician can stay to their intended interpretation. Skill alone, however, is not enough. An excellent classically trained musician might not be ideally suited to perform graphic scores, as the classical tradition rarely includes improvisational practices, or other practices that require autonomous artistic choices. In effect, classical musicians are mostly trained for close translations of traditional scores, which is a valuable skill but not one that is needed for this study. I selected the following musicians, then, on the basis of both their skill and their creative approach to music making. In addition, I chose three musicians whose instruments (flute, percussion and voice) and skills would complement each other: scores for percussion are not notated in traditional notation, jazz flutists can make music very well without scores at all, and singers are an odd one out when it comes to their instrument, as their instrument is their entire body. Keeping in mind the embodied process that is graphic score interpretation according to Casey, the voice was thought to be an interesting instrument to investigate an interpretive practice. Through the musicians' mixed set of instruments, I expect to gain insights into different approaches to performing the same scores.

- **Berend Eijkhout (BM)**[7] is a classically trained professional singer (bariton). He is most apt in performing large works, with orchestra and conductor, such as operas and requiems, where he takes the role of solist. He recognizes that such works are often surrounded by a rich performance tradition. During his professional training, he performed some graphic scores, such as *Stripsody* (1966) by Cathy Berberian and some works composed by Luciano Berio.
  Berend interprets the scores using his voice and the occasional piano accompaniment.

- **Pieter de Mast (PM)**[8] is a flutist and saxofonist, who was approached because of his activities as artistic leader of improvisational ensemble Windstreken[9]. He started his musical training when he was very young, and got into Codarts conservatory in Rotterdam, NL, to study jazz flute. He doesn't want to constrain himself to one musical activity, and enjoys teaching, composing and playing in equal measures. Pieter has no experience with performing graphic scores, but a lot of experience with improvisational music making.
  Pieter interprets the scores on flute and alto flute.

- **Baue Kunstman (BK)** just finished the bachelor programme in percussion at the Royal Conservatory in The Hague, NL. His musical training started when he was very young, and he attended the conservatory's young talent project. Baue mostly plays modern work, including work in which close contact with composers is required to get an idea of how to perform it. Baue has a lot of experience with graphic scores, abstract scores and scores that are not written for particular instrumentation.
  Baue has a more extensive setup through which he interprets the scores: a vibraphone, a set of crotales, a set of woodblocks, a cymbal and two timpani (with modulating pedals).

[7] https://www.berendeijkhout.nl/
[8] http://pieterdemast.com/
[9] https://windstreken.wordpress.com/

VII. Results

In this section, we take a look at the musicians' perception of and experience with sonifying the scores. I will present statements made by the musicians and use them to make generalized observations about their interaction with the scores. These observations will be categorized into three relevant domains:

- Musicians' reading of the scores
- Computer-generatedness and the semantic vacuum
- The role of the composer

These domains correspond for the largest part to the categories of questions I used to interview the musicians: the affordances, computer-generatedness and the role of the composer. About the first, I will repeat the fact that we are not going to be concerned with the sounding music. A few comments will be made about some musical aspects of the musicians' playing, but they will support the observations about the interpretation processes.

A. Musician's reading of the scores

First to note is that the musicians adopted mixed approaches to reading the scores. By design, any choices about how to perform were left completely up to them (outside of what is on the pages), and this resulted in different ideas about what constitutes a proper reading of a score. These ideas can be thought of as existing somewhere between a close and systematic reading on the one hand and a free performance that is merely inspired or informed by what is on the page on the other. Pieter, for instance, did not always take the specific layout of the scores and their elements all that literally, and chose to interpret scores C1 and C2 according to their holistic appearance:[10]

PM: *"With [C1 and C2], I just picked up the general idea, and I did something with that myself. I didn't try to read these from start to finish or from top to bottom. [...] If I took the effort, I could have played it from start to finish, but that would constrain me very much in telling a musical story. And that is, in the end, the goal of this."*

By contrast, Baue describes how he wanted to stay as close as possible to the scores, even though that wasn't always possible:

BK: *"The more that is happening at the same moment, the more difficult it is. The busier … ehm … this has also to do with the fact that I … ehm … try to read [the scores] perfectly. If you try to do that, from the page, and like with [C2] there are three or four things happening at the same time, so three shapes, and a line underneath, that is, at least for me, very difficult to process all at once [...]."*

Similarly, Berend recognizes the impossibility of reading the scores as close as he might have wanted, because of the practical limitations that come with the use of his voice:

BE: "*And here [C3, end of second line], I thought: damn, I am making an ending, which is bullshit because it just continues on the next line, and in this moment I thought 'well yeah', and then I started to get a little angry, first at the score and then at myself, and I thought 'but that is impossible, I have to breathe somewhere!', and then I thought 'wait, who says that I cannot breathe, that is just my own interpretation', and then I decided that I would breathe here, at the end of the line.*"

[10] All of the quotes in this section are translated from Dutch to English.

From such accounts, we can make our first generalized observation:

Observation 1: Musicians apply their own ideas and preconceptions about what constitutes a proper reading of a computer-generated graphic score.

Pieter does not read the scores literally, but uses them as inspiration for improvisation, while Baue and Berend struggle with their wish to do a close reading, but have to negotiate this wish as practical reasons prohibit it from being realized. In this sense, the strategies employed by Baue and Pieter indicate the ad hoc and practical nature of the interpretation process, as their preferred and intended strategies have to be altered on the spot for them to be able to play the scores at all.

These differences in the musicians' interpretation strategies stand out in the light of the careful framing of the scores, which was exactly the same for all three musicians. The strategies are each musician's way of making sense of the score, and they act as their toolset for making music from the scores in the absence of explicit directions or instructions. This raises the question of what other tools or guiding factors could be at play in the interpretation process.

In their approaches to reading the scores, it was observed how all three musicians used their 'musical pride' as a guiding factor. What I mean is that, while it might have been hard to quickly (within the limited time they had) and rigidly decide how close a reading had to be and how much freedom they would allow, they never lost sight of their ultimate goal: making good music. This gut feeling they share, which is presumably a result of their skill and training, was used to navigate the difficulties of having to choose what to play and how to play what was on the page. Choices that had to be made concerning how to translate certain aspects of the scores and what elements had to be left out were made with their intuitive sense of what would be the best choices *for the music*. Pieter described this most aptly in the opening quote of this section; for Pieter, the specific contents of each score are of lesser importance than their overall qualities and visual appearance, because in the end his goal is making music that sounds good and is worth listening to. Berend explains it as a particular performance mindset:

BE: *"[...] or I thought, 'now it is getting a little boring, or uninteresting, now I have to change some things for the audience, the imaginary audience. [...] I don't want the audience to feel cheated out of their money"*

I generalize into the second observation:

Observation 2: Musicians are intrinsically motivated to make good sounding music, which serves as a guiding factor when interpreting the scores.

Structural differences in the musicians' perceptions of the different scores correlate to a certain extent with the way they were designed. In section V I described how Inglis' framework was implemented to generate scores with different levels of semantic content, C1 being the most abstract and C4 being the most explicit and rich. It is here that we see the results of these differences. They are, however, not completely clear-cut. For instance, Baue described how he found C1 the hardest to interpret:

BK: *"[...] I found [C1] the most difficult to interpret, because I had to approach this one in a different way. Like we talked about before, not look at it like it is a score, but looking at the form, at the whole, and not at every individual line. Because of that it also became the hardest to play."*

He attributes the fact that he found C1 the hardest to interpret to his approach of trying a close reading of the score, like he had done with the other ones. Viewed as something to be followed closely, C1 is a highly complex and intimately structured score. In contrast, Baue recognizes that taking the score as an abstract, holistic visual, would probably render C1 the *easiest* to interpret. He also commented that he was quite unhappy with his performance of this score, because he felt like his systematic approach did not work. This opposition between a (failed) close reading and the potential of doing a holistic reading can be understood through the generated nature of the score: the visual elements are scattered in a highly random and chaotic way, while the higher-level gradients that are generated afford a reading that is easier to grasp.

Berend took a more holistic approach when interpreting C1. He describes how, for him, scores are easier to interpret when they have a balance between the amount of information and how abstract the scores are. He recognizes the complexity of C1, with its many elements, but explains how in this case this does not make this score that hard to interpret:

BE: " *[C1] has even more information, but is also a lot more abstract, so it is quite nice, because I am not forced to process all this information. I feel more freedom there.*"

The abstract nature of C1 affords Berend a free interpretation strategy, in which a close reading is not required for a valid performance. He describes how he traversed the score *"like a metro, or a platform video game"*, a strategy that combines the spacing and low-scale interplay of lines and circles with the holistic reading that Baue described.

We can see how these differences in perception that arise out of different approaches to close reading and allowing oneself interpretative freedom do correlate with the semantic levels in the different scores in some cases. Even though Berend and Baue differ in opinion about their experience with C1, they agree that this score could afford a large amount of freedom that makes interpreting it less forced or constrained than more concrete scores. However, when we look at C4, the most semantically rich score, we don't see agreement on how to deal with these meaningful elements. Berend describes how C4 is a good example of a score that balances the amount of information and the freedom that he allows himself, that is afforded by the score:

BE: *"[C4 was the easiest] because of the ratio between abstract and concrete, I think. On the one hand the amount of information that was there, so there wasn't too much information, but also not too little, and it was sufficiently concrete for me to be able to do something with it right away, without that it was constraining, that it was super concrete."*

Pieter, however, finds C4 one of the hardest to interpret because of a similar reason:

PM: *"Admittedly [In C4] there are recognizable notes in a staff, but yeah …, [C1] contains a certain idea, and C4 does not contain an idea. It is more like a cluster of notes packed together. Then you are bound to those notes, which already constrains me. [...] What to do with them, I cannot get out of the score that easily."*

For Pieter the concrete information, the notes that are in the score, is constraining. As actual musical material, the notes force Pieter into a rather small interpretation space that has to be translated

successfully into music. In contrast with C2, where he took a liberal approach of ignoring the horizontal dimension of time in the score and playing the shapes in a free, experiential way, C4 is too concrete. Of C4, Pieter performed three versions, as he wanted to do this score again to improve his handling of the musical material. In the third version, he relied heavily on a rhythmic element that he did not actually read in the score, but he felt he needed to make the piece more musically interesting.

Observation 3: Different strategies used by the musicians resulted in differing perceptions of the relative difficulty of interpreting scores that do not correspond with their level of semantic content.

Because the musicians each apply their own ideas of how to read the scores, different strategies and subsequent struggles can be seen. Although some similarities arise across musicians and scores, the most interesting characteristics seem to come into play on the individual level, and are likely the result of interpretation strategies that were adopted early in the interpretation processes of each musician for each score.

B. Computer-generatedness and the semantic vacuum

In the presentation of some practical aspects of the interpretation process above, which could be argued to apply to 'ordinary', non-generated scores as well, we have already touched upon some topics of meaning that relate to the generated nature of our scores, by focusing on musicians' reading strategies of inherently meaningless visual elements. From it, we can already see how little the musicians struggled with the computational aspect of the scores' ontology while translating the scores into music. For me, it was interesting to see how the influence of computers on the scores barely came up in the live sessions until I started to question them about it. Then, it turned out that they had a lot of thoughts about it; it just seemed to not matter that much to their experience of the scores.

Pieter, for instance, recognizes the inevitable influence of the human hand in the generation process:

PM: *"It is certain that they were put in a computer by someone, of course [...]. It is made* with *a computer, but it is made* by *a human, so it seems to me, because you have to direct a computer to get this."*

Pieter had very little idea about what strategies could have been used to generate the scores or what the signs of computer influence were, which makes his immediate intuition about the need for human influence the more telling. For Pieter, the scores could just as well have been made in a Sibelius-like program for the production of graphic scores, which means that his conception of what a computer-generated score is did not interfere all that much with his interpreting and playing the scores.

Berend could see the computer's hand in the scores. He points to the shapes in C2 that look very generated, the large snake-like line in C3 that doesn't quite look organic and the weird angles the connection lines make in C4:

BE: *" [...] it is not clear whether it is meaningful or coincidence. I think a human would have placed the lines either more straight or more at an angle, to clarify what they mean."*

About these instances, his intuition is of course correct: there has been no human revision to clear up such ambiguities. However, Berend also finds the human hand in the scores; he finds C1 suspiciously

finished and too aesthetically pleasing, to the degree that he does not believe it was made by a computer.

Concerning the actual interplay between me and the system, Berend suspects that methods of Machine Learning have been used. He sees this in the fact that coherent staffs with notes are generated. Of course, the coherent staffs were not learned through Machine Learning but implemented by me as staffs. The interesting line of inquiry that can be followed thus is that Berend does see through the simple phrasing of 'computer-generated graphic scores', attributes the human influence incoherently to different aspects of the scores, but in the end does not seem to be hindered in his performance by this fuzzy conception of the computer's part in the generation process.

Baue had a rather good intuition about how the scores were likely made. Although he opened his answer to my question with the quintessential comment

BK: *" [...] I* know *the scores are computer-generated, and I wonder whether, if you would have given me these and didn't tell me they were computer-generated, I would have known or not",*

he is sharp in describing what he thinks has been my influence and what has been the computer's. For instance, he describes the neat arrangement of the horizontal systems in C2, while also recognizing the chaotic nature of the shapes. These accounts can be summarized in the fourth and fifth observation:

Observation 4: The musicians had different ideas about what computer-generated might mean and what techniques could have been used in generating the scores.

Observation 5: These different ideas did not influence the interpretation process of the musicians in a significant way.

What is important about these observations is that they describe how the influence the musicians' ideas have on their interpretation strategies seems to be comparatively small. More specifically: even though the ideas and conceptions the three musicians had about the scores and the role of the computer were quite different, the level to which these conceptions seem to influence their approach of interpretation is not that different. Where Baue struggled with a fitting approach for C1 and felt like translating C2 was rather straightforward, Berend got overloaded by the affordances of C2 but felt most comfortable with the balance of information and concreteness in C4, and Pieter found that C4 was too constraining for him. What this summary suggests, is that intuitions or understanding about the methods of generation are not necessary for an interpretation strategy that can be followed to produce music from the scores.

Baue provides one exception to Observation 5. He describes how the absence of a composer makes our computer-generated scores somewhat different from regular graphic scores, no matter how abstract they are, because there will always be some social factors when playing pieces made by a composer:

BK: *"[...] You could say that putting musicality into [playing these scores] is more difficult here, really making a piece out of it is more difficult, because there is no explanation. And even though the interpretation is completely free, you still want to go along with the composer, or in any case, some kind of … well, what kind of person they are, what kind of pieces they write, that also makes it easier. And this is completely random."*

Baue describes how the complete absence of a composer and its explicit or implicit influence on the interpretation process can make it harder to make good music out of a graphic score. For him, there can be such a thing as too much freedom, which can make the interpretation process harder. He extends this statement to include other musicians as well, who can play a role in shaping his interpretation strategy of graphic scores by collaboration in ensembles or just through being there.

Another way of phrasing the above observations is in terms of the semantic vacuum. The way each musician has their own particular approach to each individual score, which in their cases had to be devised within minutes, reveals how the vacuum's effect on the interpretation process is limited. When they are asked to play scores, Pieter, Baue and Berend play the scores. Even though Pieter found it hard to translate C1 into music because of its abstract appearance, he allowed himself the freedom he felt he needed to make something out of it. Baue describes how he approached C2 as it being a piece of music, and plays it accordingly, even though it was impossible for him to play everything perfectly. Berend explains that he takes liberty with his breath and phrasing in C3 that is necessary for him to make something musical out of the score. All of these statements are instances of the musicians choosing to *make good music* over a literal reading of the score, even though they sometimes wanted to. I will rephrase Observation 5 to include the semantic vacuum:

Observation 5a: The existence of a semantic vacuum did not influence the musicians' interpretation process in a significant way.

C. On the role of the composer

We have seen how the semantic vacuum gets filled by the input of the interpreting musicians. Although they described some differences in approaches they took because of the absence of a composer who provided a context for interpretation, they had little difficulty playing the scores when asked to. The question that remains to be investigated is: who might be considered the composer in our context where the scores are generated by a computer? Or: could such a role be ascribed to any particular contributing party? After all, the communication and translation processes are even more muddied than already was the case with ordinary graphic scores where the sounding result is often largely 'composed' by the performer.

About this topic, we see two different opinions in our musicians. We have already seen Pieter's thoughts on the matter: the human influence and creativity overshadows the influence of the computer. Indeed, he thinks the composer is me, the system designer and programmer. For him, there is barely another answer possible. Although, he is not completely sure about this, which is a result of his lack of intuition about what the division of labor could have been. Again, his position and the fact that he nevertheless sees the scores as music without question reveals how little such nuanced questions about authorship and compositional intention mean to him in the face of music being made.

Baue agrees with his view that I am the composer of the scores, although his perspective is different. Baue makes a distinction between musical ideas and artistic ideas, the first being the written music, which codes for a sounding result, and the latter being something more general. For him, although there might be a lack of musical ideas in our computer-generated scores, there certainly is an artistic idea: music that is being generated by a computer. He describes,

BK: *"In the end, there is always someone who puts things in motion, who brings an idea to the table, and who wants to show this idea in a certain way, whether it is through music, or dance, or something visual, or a book, or theater, or whatever. [...] If I would perform these [scores] in a concert, then that would also be a performance of a musical work, with the idea 'can a computer make music?'*

Baue compares our context with the painter Jackson Pollock, and the way he tried to minimize his conscious influence on the paintings. For Baue, Pollock is just as much the painter of his works as I am the composer of these pieces, even though I outsourced some of my creative power to the computer. I did not bring *musical* ideas to the table, but I did bring another, more general idea, which for Baue is enough to start making music. He also compares our case to John Cage, who used aleatoric elements in his music. These elements rendered the sounding results largely outside of his influence. However, for Baue, this does not mean Cage is not the composer of his pieces. His idea was just to use non-traditional source material for his music.

Berend turns this idea somewhat on its head. For him, the computer is the composer. He compares our case with the case of a composer he knows who uses computer-generated source material to make their music. They algorithmically generated large amounts of melodies, and cherry-picked and rearranged these melodies to create a finished piece. In this case, for Berend, the human is the composer, even though the material was computer-generated. In our case, it is the other way around. The computer composes a particular piece out of source material that I, the human, provided.

However, Berend also recognizes the more complex nature of the role of composer. Even in older times, Berend notes, the composer was not always one person who did everything to write the music. He describes how composers sometimes wrote simple melodies and then let their students expand on those melodies to write pieces of music. Also in our case the role of the composer might be more nuanced. For Berend, parts of the traditional role of the composer get reassigned to different parties in the musical communication process:

BE: *"You could say that the role of the composer in this case has been divided, and went partly towards the person who programmed the computer, and partly to the performer. [...] 'Composer' is of course only a label that we assigned to a whole of different tasks: coming up with what it has to be about, arranging, orchestrating, etc."*

We can summarize these accounts into the last observation:

Observation 6: Musicians have different opinions concerning whether and how the composer-figure has been replaced by the generating system.

VIII. Discussion

The observations provide us with some useful empirical material through which we can reflect on the different aspects of this study. I present reflective arguments on three levels of inquiry:

1. The first level concerns the conceptual aspects of the study: the semantic vacuum, the influence of the affordances on the interpretation process, the framing of the scores and the role of a composer.
2. On the second level, I take a critical look at the theoretical implications of our observations. I will revisit the semiotic model that was built by Nattiez, Anderson and Inglis and reflect on the human influence in generative system design in the light of the discussion by Bidgoli, Kang & Llach.
3. The third level contains the conclusions we can draw from our adopted methodology in exploring the field of computer-generated graphic scores. I will discuss three choices that were made during the different stages of the study and what could be other perspectives on computer-generated graphic scores that come from choosing different options.

These levels do not correspond with the categorisation adopted in the results section. They are presented as such to group together the most important points of reflection. It is worth restating here that for a large part the study's intended contribution is to serve as a foundation for further exploration of computer-generated graphic scores. In all three sections of discussion there will be suggestions for revisions of the experimental setup that could be worth trying and for other perspectives on computer-generated graphic scores that will help understand how to study them.

A. Conceptual analysis

We've seen how the semantic vacuum, in the end, gets filled rather naturally by the musicians. This is supported by Observations 1, 2, 5 and 5a. Baue described how it does not matter that much to him that musical ideas are missing from the scores, as he approaches the scores as pieces of music anyways. Pieter latches on to what he sees on the page and uses it as inspiration to make what he thinks is music worth listening to. Berend saw the particular areas of freedom needed for him to make music out of the score emerge on the go, as it were. This makes sense, of course, if we consider that the amount of compositional meaning in the scores was kept to a minimum by design. The intentionless system did not provide the ideas the musicians need to translate the scores into music. In the absence of these ideas, they put in their own ideas. What other options do they have, except maybe to give up and refuse to play the scores, or play something arbitrary under the flag of a free interpretation?

Observation 3 ties into this line of reasoning in a slightly more nuanced way. The fact that perceived levels of difficulty did not correlate with the levels of semantic content that were implemented in the scores implies that semantic content in the scores is not a relevant criterion to consider when thinking of the semantic vacuum. Just as complex AI-techniques were considered to be out of place because of their quality as semantic veneer, so the levels of semantic content in the scores can be argued to be superficial in the light of other aspects of the interpretation process that played a role. Recall how Pieter in particular did not know what to do with the notes in C4 and how Baue struggled with the highly abstract score because it did not afford a close reading in this context. The semantic vacuum, the alleged absence of meaning in the scores, is navigated by the musicians through a combination of factors that arise in an ad hoc manner during each encounter with a score.

In part, we can understand the musicians' strategies through the scores' visual affordances. We have seen how the elements on the pages all were negotiated. Some were played while others were left out; some were taken literally, while some were loosely interpreted through their visual appearance. For the musicians, a large part of their perceived task consisted of choosing what elements to play and how to play them. Berend described a direct relationship between how difficult it was for him to interpret a score and its affordances/abstractness ratio. It is good to have the relationship between the scores' affordances and the musicians' playing reassured, as it indicates that our implementation approach of generating simple algorithmic elements was sufficient for creating potentially engaging scores.

C2 was recognized by all three musicians as a score with too much information. I could call this an affordance overload. While Baue and Pieter made similar comments on the amount of information to be processed in C2, Berend describes it most aptly at the moment the video playback of this score started:

BE: *"I immediately thought, this is too much information. It is impossible to process this. I started without knowing precisely what information I would eliminate. [...]"*

All three musicians had a different strategy of dealing with this problem. As mentioned, Pieter chose to leave out the horizontal time-axis, Berend adopted an approach of ad hoc choosing what elements to perform and what elements to leave out and Baue chose to perform *"as good as [he] could"* to play what he read on the page. Interestingly, this overload was only a problem for Baue because of the session setup. He describes how C2 would be the most promising as a piece of music, were he given the time to practice and rehearse it. His own perceived lack of precision and consistency in the performance was for him a matter of failing his musical standards.

The affordance overload in C2 needs to be reviewed from an implementation perspective. Although the second of the stopping criteria for developing the generators concerned whether or not there were sufficient affording elements in a score, an upper bound was not considered. While I explicitly opposed the adage 'more is better' in the implementation phase, I did not think C2 would suffer from affordance overload. Now, while not attempting to trivialize this problem, the ad hoc negotiation and intuitive navigation of the elements of this score by all three musicians indicates that the guiding factors utilized in their strategies, i.e. their ideas and preconceptions from Observation 1 and their intrinsic motivation for making music from Observation 2, make it quite manageable for them.

This has close ties to another factor that contributes to the musicians' lack of effort in circumventing the semantic vacuum: the way the sessions and the scores were framed. I presented the scores as computer-generated, and set up an environment in which it was natural for them to start playing after mere minutes of inspecting the score. In such an environment, in which there is very little time to think or reflect, whatever comes to mind for interpreting a score will do in that moment. This could be called a success for the framing aspect of the experimental design. The goal was to be clear and simple: what the musicians needed to know was that they were going to play computer-generated graphic scores, and nothing more. They were handed scores, which they correctly assumed to be computer-generated, and started playing them.

The success I aim at lies in the fact that this way of framing was chosen to postpone in-depth aspects of discussion on their interpretation process until *after* the actual playing, and as such it

worked. This order of business was needed to ensure an untainted perspective from them towards the scores. We mustn't forget, of course, that other ways of framing would have led to different loci of focus and areas of inquiry. Calling a certain way of framing a success can easily be an empty statement that comes with the danger of becoming a self-fulfilling prophecy. For, if we were to frame something with a certain outcome in mind, and the outcome becomes reality, it is easy but misleading to conclude that either the framing or outcome teaches us anything trustworthy, if the framing would force us into some unnatural state with no resemblance to any real-world processes. In our case, this danger was negotiated by keeping the framing as simple as possible, and giving ample space for the musicians to express their perspective before getting into areas of focus that were particularly important to our goals.

Baue's comments on the social factors at play between musicians and composers of graphic scores and the lack thereof in our live sessions are notable exceptions to the rather small influence of the semantic vacuum. It seems the effects of these factors, which are social in nature and have little semantic bearing, are rather insignificant when compared to the proactive stance the musicians take, which we've seen in Observations 1 and 2. The fact that Baue noticed the absence of a human composer, however, does raise questions about my already complex role of researcher/system developer/audience. For, if Baue's account suggests anything, it is that social factors are always at play, be they between musician and composer or between musician and any other social party that is involved. New experimental setups should be employed that address these factors, for instance where the roles of system designer and interviewer are separated, or where an audience is incorporated to change the performance dynamics, as to gain understanding of different social aspects and their effects on graphic score interpretation.

Some more words on the role of a composer are in place. Observation 6 describes how the musicians' views on what a conceptual composer should be in our context of generated scores are very different: Pieter thought the composer was me, Berend thought the composer was the computer, and Baue thought the labor was divided. That these three quintessential perspectives are all present in our group of participants is telling. In addition to the fact that it informs us about the role of a composer in our context of generated scores, it tells us that such a role in general is less well-defined than one might think. For, it is unlikely that such different views would arise if a composer was an otherwise properly understood concept. We must not forget that the surge in graphic score practice in the 1950s indicated a clear move away from the composer's power over their own music. With that, the relatively rigid role of the composer itself became increasingly unstable, as did what is called the work-concept, the finished piece of music that is created by the composer and is to be played in close detail (Magnussen, 2019). Because of these instabilities, it is impossible to conceptualize what a 'composer-system swap' would mean. With graphic score practice in the 1950s, power relations between composers and performers were already shifting; with computer-generated scores, it was to be expected that different parties contributing to an interpretation context (such as me, the musician and maybe even an audience) would share in the meaning-making process.

In section IVa I wrote how the semantic vacuum is a hypothesis of sorts, because one thing we can reasonably assume to be different between a graphic score and a computer-generated one is the lack of intention in the latter: a regular graphic score is *intended* to be a score, and thereby functions as a vehicle of meaning between the composer and the performer, while a generated one is not. While it is true that our generated scores were not intended to be musical scores by a traditional composer, they were definitely intended to be musical scores, as their raison d'etre was to be played by musicians. As even Young's absurd horizontal line composition was played by serious musicians

because it was actively presented as a musical score, it makes sense that our generated scores would be played through a kind of intention that Baue recognized as an artistic idea. After all, one horizontal line is not a nuanced work with affording elements as potential music; it is more of an experiment, or even a gimmick. One can intend all they want with such a score, it will not afford any more playing. Just as Young's piece functions as a vehicle of meaning for his experimental ideas, so do our generated scores function as vehicles of meaning that is indeed in the mutual understanding between the performer and, in this case, their justified belief in the framing in this study and that the artifact is a musical score.

B. Theoretical implications

The rather small influence of the semantic vacuum makes sense in the light of the historical context of graphic scores I sketched in section IIIa. One of the main points from that section was that composers during the mid-1900's sought ways to break the hegemonial power of the composer. This they did by outsourcing some of the artistic choices to the people performing the music. Swapping the composer with a computational system can be seen as a continuation of this trend of diminishing the power of the composer. A computer, after all, has no intention. In effect, the power of the performer becomes even greater when the composer is a computer. Based on our observations, we couldn't go as far as to say that any relinquishment of power on the compositional side of things results in an equal gain in power on the performing side, as some sort of zero-sum-game. We did, however, see how naturally liberties are taken by interpreting musicians in the face of musical pressures which they were trained to handle.

Also in the literature there are hints of the natural closing of the semantic vacuum. Naturally, graphic score scholars recognize the importance of the performers' influence on the sounding result. George Athanasopoulos (2010, p302) writes:

> A large portion of the "potential" music originates from the performer, and the less information passed from the score, the larger this portion is.

He writes this statement as justification for his hypothesis that there might be common ground in interpretation between performers that come from similar cultural backgrounds, but it can be applied to our case: in the case of computer-generated scores, where very little information or ideas are present to guide the performers' process, the portion of potential music that comes from the performers is very big. Athanasopoulos would support the view that a vacuum does not really exist, that it is merely an extreme case of absence of information, with a large portion of necessary initiative by the performers.

The semiotic model by Brian Inglis and Virginia Anderson that was described also suggests that the lack of intentional meaning might not be a big problem for the music making process. Recall that in this model, poietic as well as esthesic processes play a role. The poietic process is productive, and corresponds to the composer's composition practice as well as to the performer's performance practice. The esthesic process is receptive. In the model, Anderson rightly sees the processes between the composer and the performer as going in two directions; the esthesic process, however receptive, is an active one. This active role the performer plays in translating the (musical) ideas into music is precisely what we see happening in our live sessions: in the absence of poietic material, the musicians bring their own material and make choices that serve their conception of making good music, as we saw in sections 1, 2 and 5a.

However, following the above line of reasoning more closely, we can see how the foundations of the semiotic model come to be under pressure. As Anderson writes, “Nattiez’s model makes sense only in the strictest total serialism in which the performer complies literally and selflessly with the score” (Anderson, 2013, p134). This is the reason she feels the need to extend the model to include the performer. Indeed, Anderson recognizes that Nattiez’ model requires a rigidity of its contributing parties that does not seem to be grounded in practice and experience. If the division of power, or of “portions of potential music” is as flexible as she suggests, which is supported by our observations that interpretation is indeed a thoroughly active process, we can even question the necessity of any aspect of the poietic process of a composer. Music will be made even if this process has been deconstructed. While of course, as Pieter recognized, a creating agent is always needed to make a score, it can hardly be said that such an agent has to be engaged in a similar poietic process as was meant by Nattiez. It seems, then, that Anderson and Inglis do not go far enough in their framework, in which the active role of the performer is highlighted. To investigate this, a study could be devised, which would also rely heavily on a nuanced framing aspect, that centralizes the question of whether a visual artifact has to be presented as a musical score, i.e. as something that is *intended* to be a score, in order for a musician to read music out of it.

Some of the most interesting discussions during the interviews came up when I asked about my and the system’s influences on the resulting scores. While we saw three categorically different views (Baue pictured a design vs. randomness division, Berend suspected Machines Learning techniques were used and Pieter thought it was mostly me who made the scores), it is important to note that all three recognized that I played at least some role. Computer-generated, while the scores were phrased and framed exactly like this, is a phrase that hides its nature behind a mask of simplicity.

As described in Observation 4 and 5, the musicians’ different views on the matter did not seem to have a significant impact on their interpretation strategies. That they all recognize traces of my influences is, however, a problem from a theoretical standpoint. For, if crucial affording elements in the scores were actually the result of human design, instead of the algorithmic contributions of the system, questions could be raised about the computer-generated nature of the scores. Let’s consider C1, where the visual elements were interpreted by the musicians as forming a big gradient: light to dark, empty to full, small to big. While the specifics of this gradient were not actively conceptualized as something to be turned into music, the gradient itself was, in large parts, the result of design. All the elements that form the gradient were changed by a parameter that was defined as a function of the y-position of these respective elements, and it was consciously programmed like this. The fact that all three musicians recognized the gradient as the central element of the score (and as the central element to be transformed into music) could be called exposure of what nuances could be at play in the generation process and the deceptive simplicity of the phrase computer-generated.

To the musicians, this exposure means different things. As we’ve seen, Pieter saw mostly my influences when considering the scores. It is possible that for him, the fact that I consciously programmed a gradient that he immediately recognized as something to be played places me in a position where a poietic process could be ascribed to me. He only started considering other options when I pointed out to him that they might exist. Baue, on the other hand, made an explicit distinction between musical ideas and artistic ideas. His account suggests that while I may be the author of the latter, the former were rather undefined, which is where he came in. It is not precisely clear what categories of ideas are the subjects of a poietic process in Nattiez’ model, or whether the two require their own processes that need to be considered. I would say that, whereas Nattiez was writing on the semiotics of music, the more general artistic ideas were likely not his primary concern.

I am unable to provide definitive solutions to these problems. In computational creativity research, this question of how to assess different influences on the generated artifact is just now surfacing and the present thesis is merely an attempt to complement this discourse. In this study, the scores have been intentionally framed as simply as possible, as to eliminate any impactful effects of the framing on the musicians' perception of the scores. Serious objections by the musicians about the generated nature of the scores would have been stimulated and welcomed as precious material for reflection. Trying to conceal the 'socio-technical nature of AI systems' (Bidgoli, Kang & Llach, 2019, p1) is a way of working around this problem without acknowledging it, endangering any results that rely on the assumption that the artifacts are, indeed, computer-generated. If anything, the present study shows that the field of computational creativity would do well to start acknowledging the crucial human influence on their computational systems, so as to not overestimate their claims on the alleged creativity of these systems.

More specifically, the line of inquiry exhibited in this thesis could benefit from an adaptation of the system design and implementation stages that centralizes a concern for human influence in system design and machinic surrogacy. The example of the gradient in C1 above is but one of the aspects of the scores for which a similar argument could be made, for instance in the coherent systematic layout and 'H' and 'L' symbols in C2. Furthermore, objections could be raised about the stopping criteria that I used to determine when a score is finished. Especially criterion 3, that includes the word 'feels', would be hard to defend against an argument that is inspired by the socio-technical nature of AI systems.

C. Reflections on methodology

Because many of the avenues explored in this thesis have yet to be paved, reflecting on choices concerning what aspects to include in the research and how to approach these aspects has become part of the methodology. These choices were often based on gut-feeling and intuition about what areas of inquiry would be potentially the most fruitful to study. The necessary injection of ethnographic aspects to the methodology of the study is the embodiment of this. Allowing oneself the freedom to explore different perspectives on the subjects and problems at hand has the advantage of being able to find out where the heart of these subjects and problems lie, and how to study them in the most natural way. On the many explored routes in all the different aspects of this study, I will not dedicate too many words, as the space is required for more poignant discourse. The point is that while trying to get a firm perspective from which to study what were considered the most interesting aspects of computer-generated graphic scores, some choices were made that seemed more deciding than others. In this third stage of reflection, I will focus on three of such choices, discussing why they are important and what other options might yield for future research:

- Focusing on the score-performer relationship
- Adopting the 'first encounter' interpretation approach
- Not adding generated instructions to the scores

The focus of the study is the interpretation process of musicians. To investigate this, it was deemed best to leave out the attendance of any kind of audience. When an audience is present, the process becomes, in addition to being interpretation, a performance, which comes with additional musicological and social aspects to an already complex area of interest. In addition, bringing the aesthetic value of the music into the equation might obscure the focus on the musicians' perception of the scores by coupling certain performances of scores with judgment on beauty. For these reasons, the choice was deliberate and sound.

However, during the live sessions and their many playbacks, it became clear that it is somewhat unnatural to consider the meaning of scores without some assessment of what music is made through them. The quality of music was a primary driving factor in the interpretation strategies of the performers, and this resulted in them doing 'what is necessary' to achieve such quality. This could mean following the score closely, but also leaving aspects or dimensions of the scores out of the translation. Musicians use what is there (and not there) on the page to be guided or inspired towards music that is interesting to them, and ultimately, so it is hoped, to an audience. They are open to reading musical qualities that are not necessarily written explicitly, such as harmony, rhythm and tension curves. It is in this way musicians engage with the scores, extract from them what they need to transform them into experiential music.

In addition, the framing of the scores to the musicians was intended to yield no confusion about the fundamental ontology of the scores: they are musical scores, and should be understood as such. This was designed as such to postpone discussion on this matter until after the performances. Retrospectively, it is possible that such strict framing might not have been necessary, as there was no reason for the musicians to doubt or object to the generated artifacts being musical scores. In any case, the fact that the scores' neutral framing nonetheless highlighted their ontology as potential music, reveals how the scores can not really be seen as separate from the resulting music.

Focusing on the generated scores as potential music will open up ways of understanding them that can incorporate some form of qualitative judgment. As we have seen that generating graphic scores with a computational system does not impede the musical communication process to a significant degree, a sensible next step would be to consider what constitutes a successful communication process, one that results in music of high quality. Methods similar to the ones employed in this study can be used to investigate the perception of musicians on their own playing, for instance. Otherwise, an audience can be incorporated into the experimental setup, which would be a sensible option if one wants to investigate the whole musical communication process, from composer through performer to audience.

Furthermore, what had great influence on the interpretation strategies of the musicians, and what might have resulted in their music-oriented approaches, is the fact that they got so little preparation time. All three commented on how they would likely have done things differently if they got the chance to look more closely at the score, devise some set of rules for it, or otherwise could have prepared their performances better. In the face of such time restrictions, it is natural they resorted to 'quick and dirty' solutions for ambiguities and doubts that arose during the interpretation process, While quick and dirty is phrased a little negatively here, as it is certainly allowed to solve ambiguities in any way that either one of the performers did, it is clear that a more critical and appreciative reading of the scores would lead to very different results. In addition, as one of the musicians' main concern was making good music, it would be fair to provide them with space to utilize their skills to this end to the fullest extent.

The paradigm that was followed to get to this setup with limited preparation time, which was chosen most prominently to make the interpretation process more transparent and accessible, is inspired by a particular compositional approach, which was described in section VI. This approach, in which first-encounter performances with scores are common and often encouraged, suited the need for an experimental setup that focuses on the interpretation process of musicians. Scores of musical works that employ this approach, such as Metaphysics of Notation, often have extended vocabularies of visual elements and are rich and complex in design, presumably to employ the potential visual affordances to accommodate the lack of preparation time.

Other compositional approaches, that do include preparation time, could be inspiration for other research perspectives. Examples are John Cage's Fontana Mix, a score of which the final, playable form has to be constructed by the performer out of some pieces of physical material and a set of rules, or Cardew's Treatise, a large, intricate work that generally benefits from at least some acquaintance with the score. For such scores, a different experimental setup should be devised, in which the performer has ample time to be with a score and think of a way of approaching and translating it. Such an approach could be combined with a more music-oriented perspective to highlight aspects of the musical communication process that were left out of this study. Different implementation approaches could also be a fruitful approach: many AI techniques, such as grammars, evolutionary algorithms and fractals, are perfect for creating larger intricate visual effects that are suited for closer inspection by musicians. Such approaches could be combined with the line of reasoning proposed by Rob Casey to investigate the cognitive structures musicians impose on the visual elements that were not designed to be anything.

While some major graphic works, such as Treatise or December 1952, leave the interpreter completely free without any written clarification, these works are comparatively rare. It seems Browne's piece, while also being quite an early work, satisfied most composers' needs to create a truly indeterminate piece, as abstract and with as little semantic content as possible. The majority of graphic works, then, comes with some written instructions. While, of course, the nature of these instructions and the ways they relate to the score vary wildly between works, it can be generally stated that score and instructions complement each other, in the sense that one is useless without the other, like a lock and its key. Works where the instructions do not refer to any aspect of the score, for instance, are unknown to me and likely quite rare.

The inclusion of some set of instructions or rules through which musicians can make sense of the scores will likely have a significant impact on the dynamics of the interpretation process, and the relationships between musician and score. Natural language, after all, is arguably human kind's most semantically rich and complex method of communication. The instructions themselves would have to be a distinct area of focus, as would the relationship between the score and the instructions: to what extent should the visual elements of the score be 'explained', or 'clarified' by the instructions? Should the instructions be perfectly complementary to the scores, or is some fuzziness allowable or even desirable? Such questions, in addition to them making the generation process more complex, make us recall the 'machinic surrogate' discussion that we described throughout this text: the score generator, the instruction generator and the relationship between the two all come with their own issues of human influence and intentional surrogacy. Generating instructions to complement a score would require an implementation strategy that takes these issues into account.

That said, written instructions could make a graphic score a more effective source of potential music. As we have seen, musicians will play the score if they are asked to. Such an open mind, that is required to engage with graphic scores in the first place, will likely result in successful translations of score to music. Regardless of whether or not the instructions actually make perfect sense in the context of the score, relationships between score and instructions will be conceptualized by the musicians that are either straightforward, creative, hilarious or surprisingly affording, all of which are perfectly good sources for making music.

IX. Conclusion

In this thesis we have taken some first steps into research on computer-generated graphic scores. We shed light on the interpretation strategies of three musicians during live sessions and their conception of the semantic vacuum while sonifying scores that were generated by a computational system. Based on observations made from qualitative documentation methods we can assert the careful conclusion that the influence of the semantic vacuum is limited. In some aspects of the interpretation process the musicians recognized some effects of the semantic vacuum, or the absence of a composer-figure, such as the fact that the presence or even the existence of a composer in an interpretation context inherently comes with some social influence on the process. Overall, however, our musicians showed to be able with relative ease to make music out of the scores. They negotiated the aspects and elements of each score with strategies that responded to ad hoc problems and opportunities that arose right before and during playing. A combination of factors contributes to this conclusion: the fact that the scores were framed as musical scores beyond questioning, the visual affordances of the generated scores and the musicians' motivation to make good quality music are among the most important contributors to this smooth interpretation process.

The power that was assigned to the semantic vacuum in this study betrays some naiveté, in the sense that its conception is somewhat simplistic in the face of theoretical and historical understanding. Writing on graphic scores describes how they are used to play into the necessity for interpretative action by the musician, and diminish the composer's influence on the sounding result; this necessity is in the perfect position to fill the vacuum quite naturally. In a sense, expecting a clean composer-system swap could be argued to adhere to an out-of-place modernist perspective, as it fails to recognize the fluid nature of musical meaning making and interpretative practice.

Then again, what this thesis is intended to contribute is in large part to provide some stepping stones one can use to study the ontology of computer-generated graphic scores in more detail. To that end, the areas of focus have been subject to extensive deliberation over the course of this study, for the reason that taking some first steps can only be done once and they hence have to be taken with care. Naturally, many other perspectives deserve a similar manner of attention. In addition to presenting the preliminary empirical findings described in this text, contributions of this thesis then include the reflections on suitable methodologies to study computer-generated graphic scores and what aspects are the most promising areas to research. The ethnographic and qualitative documentation methods employed in this study reflect the complex and experiential nature of the material at hand. We observed a real-world series of processes in which trained musicians are asked to make music, on the spot, from alien and mystically framed musical scores. In my view, although they are phrased like this partly as a result of the interplay between conceptual development of the study and experimental design, such processes can only be understood when the spheres of influence are properly respected and negotiated.

Some suggestions for further research have been discussed. Most notably, it was shown how the incorporation of the sounding music as an area of focus, the increase of preparation time for the musicians, and the inclusion of generated instructions to the scores are relevant and potentially fruitful topics to explore. These three suggestions embody what are deemed the most critical choices that were made concerning what this research was to include. In this study, options were soundly chosen that simplified the area of inquiry. With careful consideration, however, the investigation of computer-generated graphic scores can benefit from alternative approaches concerning the three topics: quality of played music was observed to be a primary driving force for interpreting musicians

and so it would make sense to consider it as an important object of study; providing more space for rehearsal and closer inspection of the scores would better fit professional musicians' training, their set of qualities and traditional composition contexts; as scores without any accompanying instructions are rare, including them will bring a subsequent study closer to traditional graphic score practice.

In addition, the thesis argued for the importance of recognizing the 'socio-technical' influences of human agents on generative practices in computational creativity contexts. The phrase 'computer-generated' may be misleading in its apparent simplicity, write Bidgoli, Kang & Llach (2019), as generative systems may unwillingly be used as intentional surrogates by their human designers and operators. In our case of computer-generated graphic scores, consequences of such surrogacy would be substantial. If it can be argued that the scores were, to a significant degree, composed by me, the programmer of the system, then a genuine vacuum will not have been created in this study. Even though, in large part through careful framing of the scores, the musicians did not question the generated nature of the scores in this study, reflection on the employed implementation approach suggests that obvious traces of human design can be found in the scores, and were indeed found by the musicians. This issue can be investigated through the employment of other implementation perspectives to generative graphic scores, which should take into consideration the dynamics between conceptions of autonomy in computational systems and influences of their human designers.

We can summarize work done in this thesis as heavily humanities-inspired work in computational creativity. Whereas this field is often concerned with the judgment of creativity of computational systems and what algorithms can be used to evoke such creativity, this thesis proposed a perspective on computational creativity that centers the embedded nature of generated artifacts: how do people interact with them, given the understanding that they are computer-generated? While not intending to trivialize the former questions of creativity in computational systems, I will argue that taking a broad perspective on generative systems and the real-world relationships they are part of will help us understand more of creativity's elusive aspects, that are poorly understood even in humans. Creativity, after all, occurs and is judged by people who have internal drives and motivations and who are culturally and socially embedded. This thesis, which fell for some traps of not acknowledging the complexity of embedded computational creativity artifacts while trying to avoid others, is testimony to the difficulties of post-disciplinary research practice as well as its potential for better understanding of complex topics in computer science and humanities.

X. Bibliography

Adhitya, S., & Kuuskankare, M. (2012). Composing Graphic Scores and Sonifying Visual Music with the SUM Tool. *Proceedings of the 9th Sound and Music Computing Conference,* 171-176

Anderson, V. (2013). The Beginning of Happiness: Approaching Scores in Graphic and Text Notation. In P. de Assis, W. Brooks & K. Coessens. *Sound and Score: Essays on Sound, Score and Notation* (pp. 130-142). Leuven: Leuven University Press

Athanasopoulos, G. (2010). Semiotic Approaches to the Representation of Music. *Before and After Music*, 299-308

Bidgoli, A., Kang, E. & Llach, D. C. (2019). Machinic Surrogates: Human-Machine Relationships in Computational Creativity. *25th International Symposium on Electronic Art,* arXiv:1908.01133

Brown, E. (2008). On December 1952. *American Music, 26*(1)*,* 1-12

Carnovalini, F., & Rodà, A. (2020). Computational creativity and music generation systems: An introduction to the state of the art. *Frontiers in Artificial Intelligence*(3).

Casey, R. (2015). Developing a Phenomenological Approach to Music Notation. *Organised Sound, 20*(2), 160-170

Cole, H. (1974). *Sounds and Signs: Aspects of Musical Notation.* London: Oxford University Press

Colton, S., Halskov, J., Ventura, D., Gouldstone, I., Cook, M., Pérez-Ferrer, B. (2015). The Painting Fool Sees! New Projects with the Automated Painter. *Proceedings of the 6th International Conference on Computational Creativity,* 189-196

Colton, S., & Wiggins, G. A. (2012). Computational creativity: The final frontier?. *ECAI*(12), 21-26.

Eco, U. (1989). *The Open Work.* Cambridge, Massachusetts: Harvard University Press

Galanter, P. (2009). Thoughts on Computational Creativity. In *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Gibson, J. J. (2014). *The ecological approach to visual perception: classic edition*. Psychology Press.

Inglis, B. (2015). Towards an Analytical Framework for Graphic Scores, and a Proposed Typology. *Putting the Graphic in Music - Notation, Analysis and Performance*, http://eprints.mdx.ac.uk/25905/

Magnussen, T. (2019). *Sonic Writing: Technologies of Material, Symbolic and Signal Inscriptions.* New York: Bloomsbury

Malafouris, L. (2008). At the Potter's Wheel: An Argument for Material Agency. In C. Knappett & L. Malafouris. *Material Agency* (pp. 19-36). Boston: Springer

Mattinson, H. & Sarkar, A. (2020). Excello: Exploring Spreadsheets for Music Composition. *Proceedings of the International Conference on New Interfaces for Musical Expression*, 11–16

Monelle, R. (1991). Music and the Peircian Trichotomies. *International Review of the Aesthetics and Sociology of Music, 22*(1), 99-108

Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, *17*(1), 67-99.

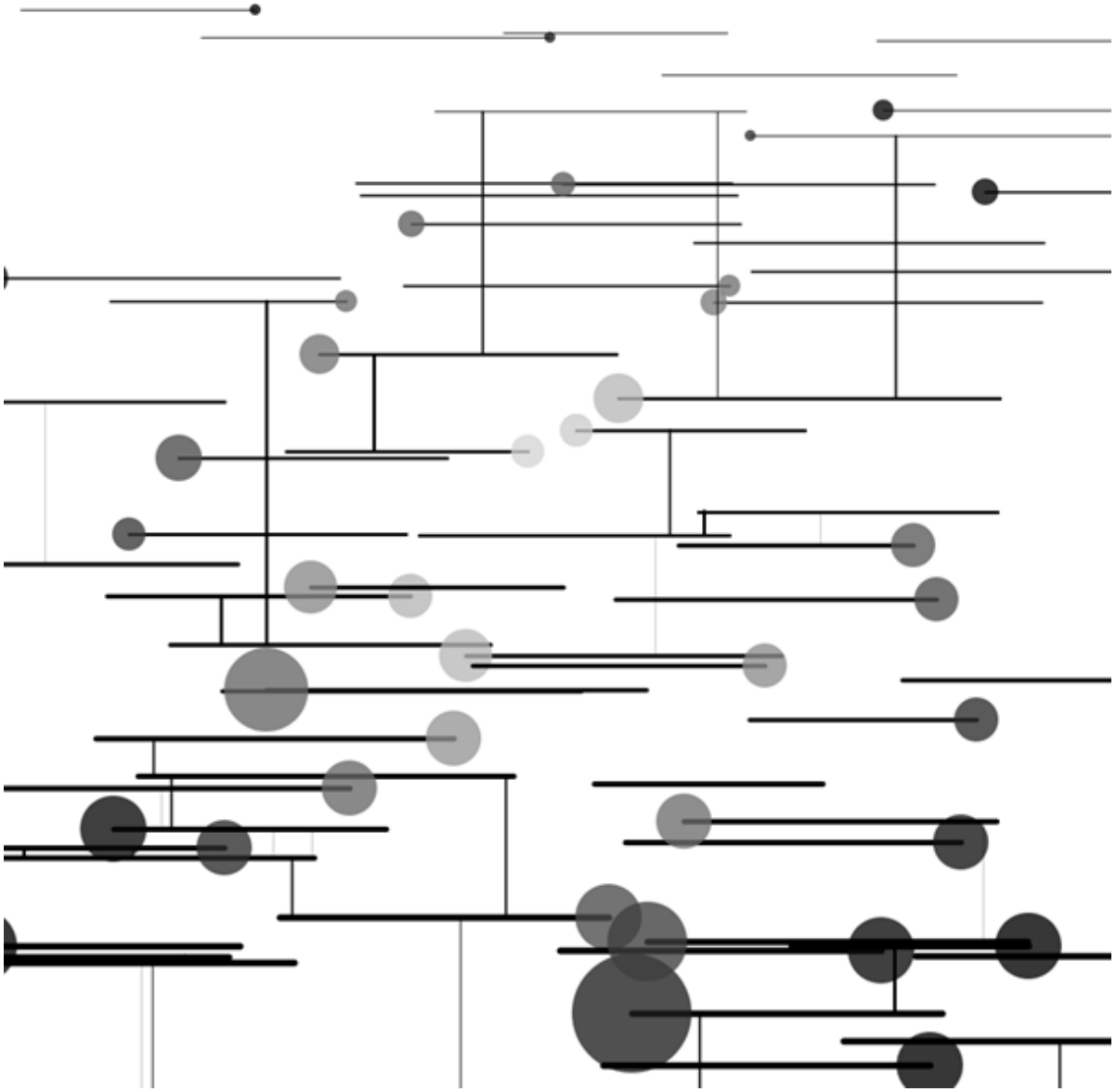
Tarasti, E. (2002). *Signs of Music.* Berlin: Mouton de Gruyter

Treitler, L. (1982). The Early History of Music Writing in the West. *Journal of the American Musicological Society, 35*(2)*,* 237-279

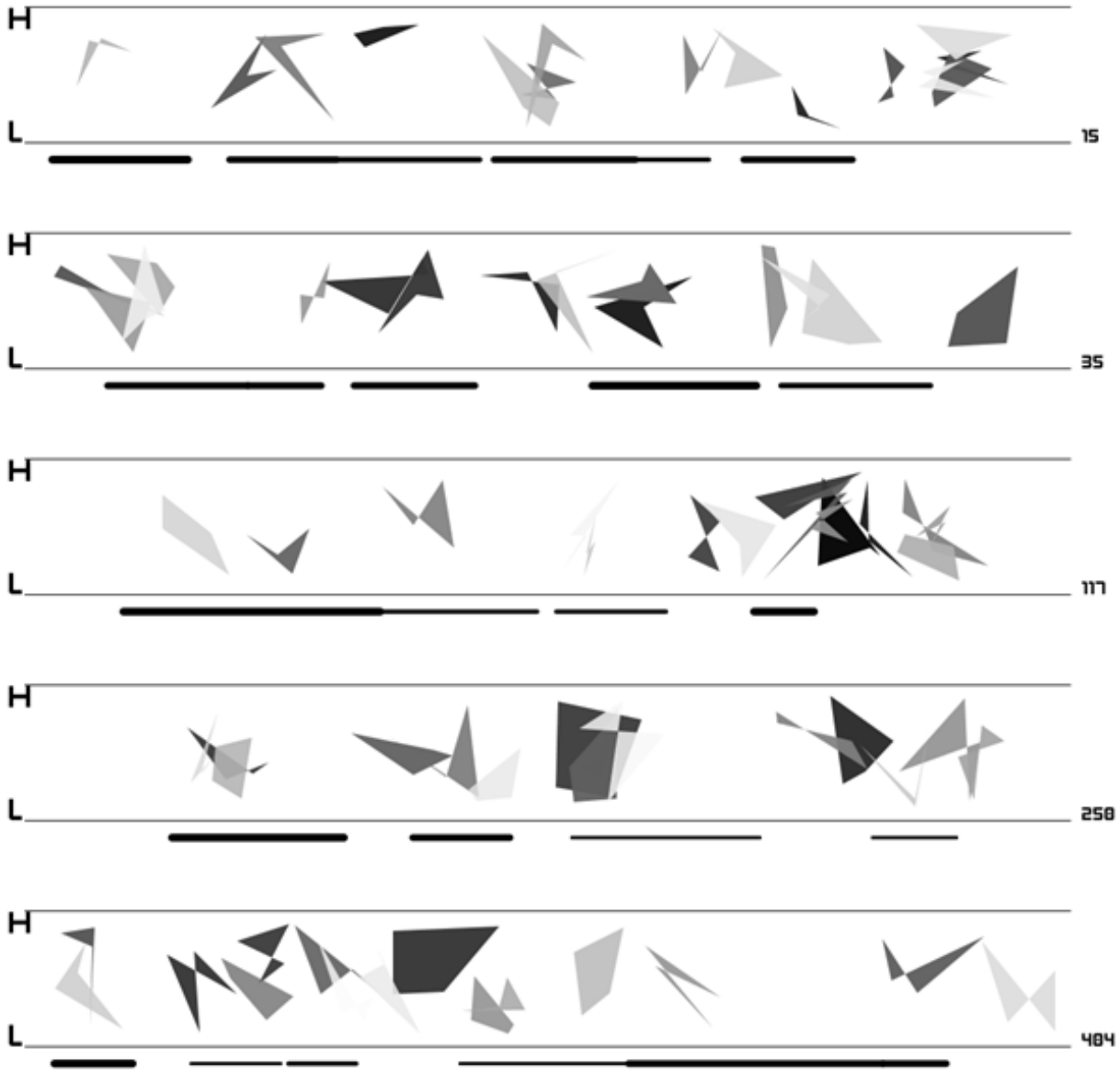
XI. Appendices

A. Generated scores

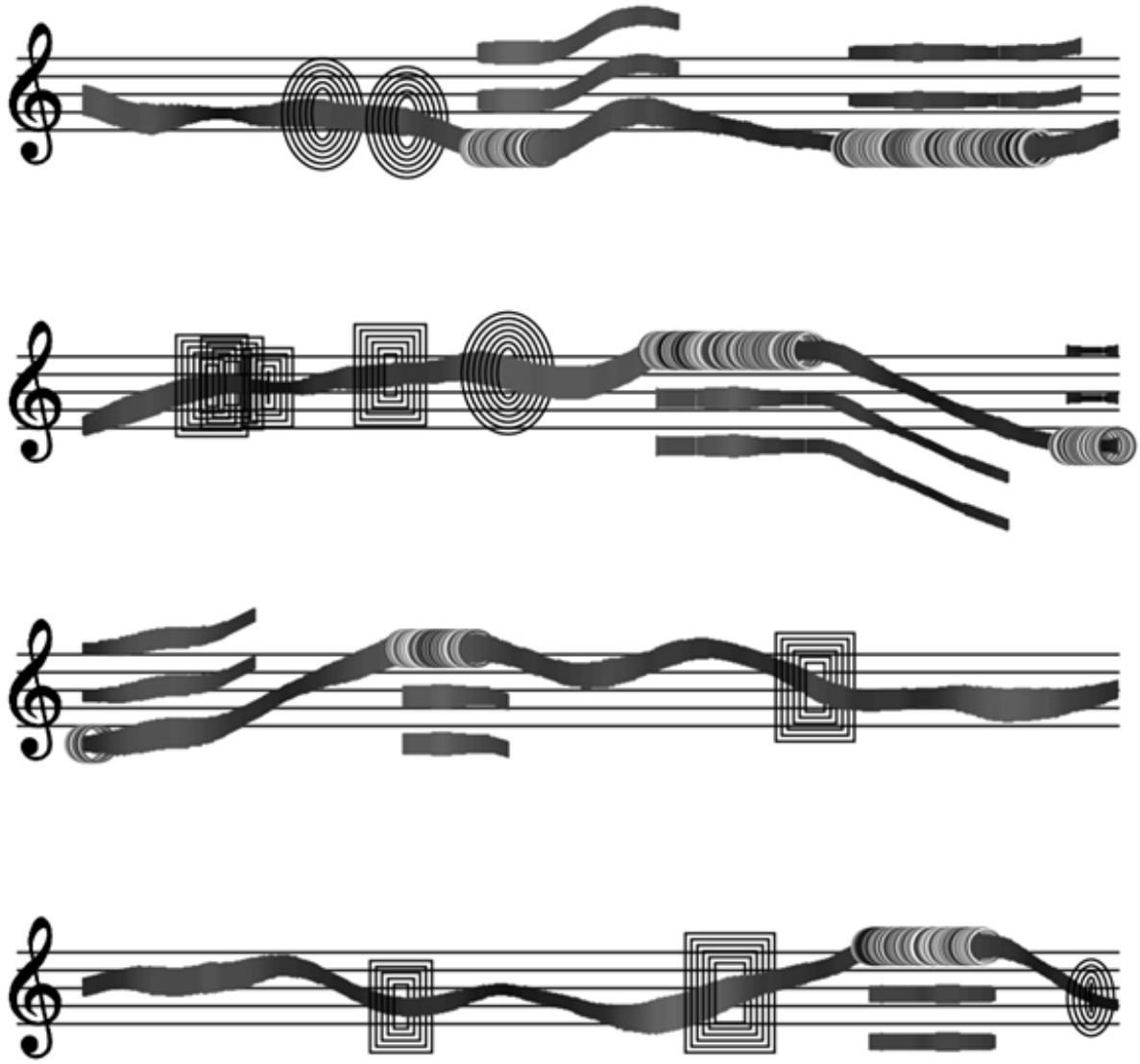
Category 1 - No musical signs, no axes



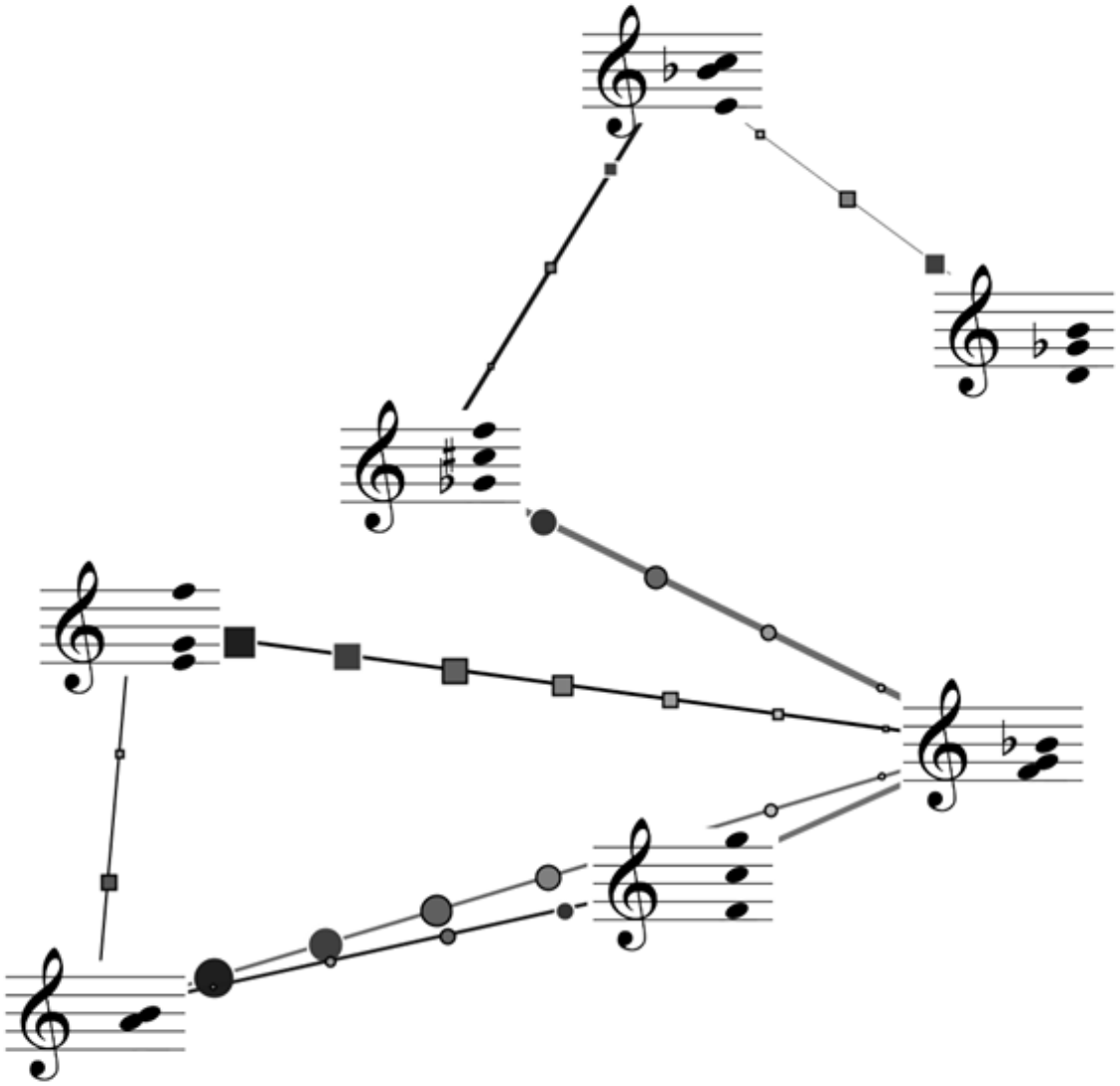
Category 2 - No musical signs, with axes



Category 3 - Musical signs, no usual symbolic meanings



Category 4 - Musical signs, with usual symbolic meanings



B. Referenced scores

Applebaum, M. (2008). *Metaphysics of Notation* [musical composition]. Retrieved from https://web.stanford.edu/~applemk/portfolio-works-metaphysics-of-notation.html

Berberian, C. (1966). *Stripsody* [musical composition]. Retrieved from https://www.houstonpublicmedia.org/berberian-stripsody-pg4/

Brown, E. (1952). *December 1952* [musical composition]. Retrieved from https://umsnjip.ch/december1952.htm

Cage, J. (1958). *Fontana Mix* [musical composition]. Retrieved from https://johncage.org/pp/John-Cage-Work-Detail.cfm?work_ID=79

Cardew, C. (1963-1967). *Treatise* [musical composition]. Retrieved from https://en.wikipedia.org/wiki/Treatise_(music)

Young, L, M. (1960). *Compositions 1960* [musical composition]. Retrieved from https://en.wikipedia.org/wiki/Compositions_1960

Young, L, M. (1960). *Compositions 1960 #9* [musical composition]. Retrieved from https://www.moma.org/collection/works/127630

Young, L. M. (1960) *Piano Piece for David Tudor #1* [musical composition]. Retrieved from https://www.moma.org/collection/works/127638

C. Interview protocols

1. Musical background interview

Can you tell a little bit about yourself as a musician?
→ Any classical or jazz training?
→ Theorist or practitioner?

Do you have any experience playing indeterminate scores?
→ If yes, what?
→ If no, any expectations for today?
→ Thoughts about music generation systems?
→ Thoughts about musical translation?

How did you choose what instrument to bring?
→ Was it deliberate?

Have you prepared in any way for today's session?

2. Semi-structured interview

a. Affordances

Which scores were easy for you to play and which ones were hard?
→ Why were these scores easier/harder than others?

How did you decide what to play on your instrument?
→ In what (parts of) the scores was it easier for you to decide?
→ Did you ever have the feeling there was more than one option for you to play?

Which aspects (parts, elements, sections, aesthetics) of the scores were easy for you to play and which ones hard?
→ What makes it that these aspects were easier/harder?

What parts of your interpretation do you think were more consistent and what parts not?

b. Generatedness

Are there parts of the scores that were indicative of the generated nature of the scores?
→ What parts?
→ What makes it that these parts look generated?
→ Were there parts that looked like they were *not* generated?

Did the fact that the scores are computer-generated influence you in your performance?

→ Why (not)?
→ How?

Did you manage to give meaningful interpretations of the scores?
→ Was it hard to make them meaningful?
→ Did you find there was meaning to be found in the scores?
→ Is this different in traditional scores?

What did you think about the musicality of the scores?
→ And what about the musicality of your performances?

c. Composer role

Who do you think is the composer of these pieces?
→ For each possible answer component: what makes you say that?
→ Is this surprising when compared to a traditional composing context?
→ Is there a composer?
→ what is the role of the composer in traditional contexts?
→ Did these scores satisfy that role?
→ What is the role of the composer in this context?

Do you think these pieces are music?
→ Did the computer-generatedness change your perception on this topic?
→ Is this different from traditional scores?
→ Was it the score or the framing?