An Automated Digital Analysis of Depictions of Child Maltreatment in Ancient Roman Writings

Author: Alexander Browne (email: alexlawford@icloud.com)
Supervisors: Dr. Cecilia Melder and Håkan Bengtsson, Uppsala University

Autumn 2019
Abstract

Historians, mostly engaging with written evidence, have argued that the Christianisation of the Roman Empire resulted in changes in both attitudes and behaviour towards children, resulting in a decrease in their maltreatment by society. I begin with a working hypothesis that this attitude-change was real and resulted in a reduction in the maltreatment of children; and that this reduction in maltreatment is evident in the literature. The approach to investigating this hypothesis belongs to the emerging field of digital humanities: by using programming techniques developed in the field of sentiment analysis, I create two sentiment-analysis like tools, one a lexicon-based approach, the other an application of a naive bayes machine learning approach. The latter is favoured as more accurate. The tool is used to automatically tag sentences, extracted from a corpus of texts written between 100 B.C and 600 A.D, that mention children, as to whether the sentences feature the maltreatment of children or not. The results are then quantitively analysed with reference to the year in which the text was written, with no statistically significant result found. However, the high accuracy of the tool in tagging sentences, at above 88%, suggests that similar tools may be able to play an important role, alongside traditional research techniques, in historical and social-science research in the future.
## Contents

1. Introduction 4
   1.1 Overview 4
   1.2 Purpose and Objective 5
   1.3 Limitations 6
   1.4 Survey of the Field 7
      1.4.1 Child maltreatment in ancient Rome 7
      1.4.2 Early christianity and the child 8
      1.4.3 sentiment analysis 8
      1.4.4 the application of sentiment analysis to novel domains 9
      1.4.5 Sentiment analysis and works in translation 10
2. Historical and Technical Theory 10
   2.1 Defining Child Maltreatment 10
   2.2 The History of Child Maltreatment 11
      2.2.1 Historical Developments in the Treatment of, and Attitudes towards Children 11
      2.2.2 Links Between Christianity and Changes in the Treatment of Children 13
   2.3 Sentiment Analysis and Similar Tools 17
      2.3.1 approaches 17
      2.3.2 interpretation of results 18
   2.4 Study Design 19
3. Material and Method 20
   3.1 Corpus 20
   3.2 Identifying Relevant Sentences 20
   3.4 Developing an Algorithm 20
      3.4.1 Simple tagging 20
      3.4.1 Naive Bayes 23
4. Results and Analysis 27
   4.1 Results 27
      4.1.1 By Decade 27
      4.1.2 By Century 27
   4.2 Analysis 28
   4.3 Conclusions 29
5. Discussion 31
References 35
1. Introduction

1.1 Overview

In the ancient Graeco-Roman world, children were treated in ways that most modern individuals would find appalling. The exposure of infants, which very often resulted in death, was widespread (Harris, 1994). Children of all ages were made subject to what would today be termed sexual abuse (Harper, 2013). Weak and deformed infants were likely to be killed, often by drowning (Bakke, 2005). The Romans made no provision for the care of orphaned or abandoned children (Miller, 2003). While both Stoic and Jewish writers expressed condemnation for these kinds of practices (Harper, 2013; Bakke, 2005), it is in the literature written by Christians that we find the largest wealth of texts voicing disapproval for exposure (Harris, 1994), infanticide (Bakke, 2005), and the sexual abuse (Harper, 2013) of children. A number of studies have argued that Christian influence changed societal attitudes towards children in antiquity (for example Strange, 2004; Grubbs, 2013; and Bakke, 2005), and the first Christian emperors certainly passed laws that were first aimed to reduce and then finally outlawed the exposure of children (Harris, 1994) and provided legal protection against the sexual abuse of slaves, including children (Harper, 2013). In the following centuries, Christian emperors invested in institutional care for orphans and foundlings, expressing explicitly Christian motivations for doing so (Miller, 2003). This study, then, aims to find evidence of a change in the plight of children, coinciding with Christianity’s transition from being the religion of a persecuted minority to dominance within the Roman word from the Edict of Milan (313) onwards.

Given this aim, it is necessary to turn to literature written at the time, as the majority of the evidence for the existence and extent of exposure and infanticide in the Roman Empire, as two important examples of children’s maltreatment, is literary (Grubbs, 2013).

I have long had an interest in the history of the era in question, in particular the approximately three century period between the end of the Diocletianic Persecution through to the death of Pope Gregory I (604), and in reading original writings from this point in history and comparing them with works from the few centuries prior, I formed the impression that there were fewer depictions of children maltreatment in the later period. I suspected this to be indicative of the growing influence of Christian values after the time of Constantine the Great, for the reasons outlined above.

Another, normally unrelated, interest of mine is computer programming: I have been involved with computer programming for over two decades and use programming tools as

---

1 Understanding the term to be defined as any human below the age of 18 years old, in accordance with the UN Convention on the Rights of the Child.
part of my work. Within the field I have read widely on the subjects of natural language processing (NLP) and specifically sentiment analysis (SA). It struck me that I could use approaches developed within this domain to research the validity of my impressions concerning the depictions of children’s maltreatment in the ancient literature.

SA uses programmed algorithms to automatically analyse opinions and sentiments expressed in texts (Pang & Lee, 2008). Existing SA research is principally focused on social media and news; accordingly, up until fairly recently there has been little attention given to the application of such tools to the study of history (Sprugnoli, 2015). There is no inherent reason why SA and similar tools should be restricted to the domains of social media and news, however (Mohammad, 2011).

In order to apply technologies developed within the field of SA to the research of the plight of children in antiquity, then, I develop an SA-like algorithm to automatically tag sentences, extracted from a corpus of texts written between 100 B.C and 600 A.D, that mention children. Sentences are tagged by the algorithm as either featuring maltreatment or not, and the results are then quantitively analysed with reference to the year in which the text was written. I explore two possible approaches to developing an appropriate algorithm; one algorithm uses a simple averaging of polarity values on a word-to-word basis, the other an application of a Naive Bayes probabilistic classifier. Results shed light on both the presence of trends in the depiction of children’s maltreatment within the texts over time, and also the usefulness of SA and SA-like tools in the analysis of large historical corpora.

1.2 Purpose and Objective

My working hypothesis is that social attitudes towards children in the Roman Empire changed with the beginning of the Christian Era; that this attitude-change resulted in a reduction in the maltreatment of children; and that this reduction in maltreatment is evident in the literature. This research aims to find if such a change is, in fact, to be found in the literature and, if it is present, to explore the nature of this change so as to better understand how the frequency and type of child maltreatment may have been affected by the Christianisation of the Roman Empire. By practical application of an SA-like tool to the study of this subject, I also aim to better understand the applicability of such tools to the analysis of historical corpora of this kind.

The research questions are:

1. Was there a reduction or change in the maltreatment of children coincident with the beginning of the Christian Era, and in what ways did this change manifest?

---

2 Herein defined as beginning in 313, with the Edict of Milan.
2. Is this change reflected in the (extant) writings produced over this period, and if so in what way?
3. Are SA-like tools useful for the analysis of historical corpora, and what can they contribute?

1.3 Limitations

My very limited abilities in both Ancient Greek and Latin dictate that I use works in translation for conducting this research. For example, it is obvious that manually tagging a training corpus, as further discussed below, requires that the human tagger understand the sentences they are tagging. This adds an additional layer of interpretation between the original text and its analysis; as Balahur & Turchi (2014) point out, SA performance can be affected by issues with translation. The same authors, however, show that SA of machine-translated texts can produce good results, so there is no reason to believe that SA-like analysis of human translated texts should not be at least equally successful. Future research conducted in cooperation with experts in ancient languages could expand the work presented in this study to analyse the original texts directly.

The corpus used, as discussed below, is relatively small, containing only 418 texts from the period in question. This lack of scale is especially apparent when this corpus is compared to those often addressed by SA, which can include millions of words (See, for example, Pak & Paroubek, 2010). As a result, this study is somewhat underpowered and so obtaining statistically significant is difficult. By addressing original-language works in the future, a larger corpus could be addressed. The manually-tagged corpus, used for training the Naive Bayes probabilistic classifier, is also small, at only 1000 sentences. This puts some constraints on the algorithm design, which in turn can affect its accuracy; nonetheless, the accuracy is shown to be fairly high. The corpus groups together different genres, including history, courtroom orations, theology, etc., the implications of which will be further discussed below. Lastly, the texts are dated based on the date when they were completed; this choice represents a compromise which is, in my opinion, appropriate for the corpus as a whole, but in the case of some texts on an individual level may be a limitation: some texts, for example, mostly consist of collations of much older writings or, while being written in a given year, may be historical works which focus on a much earlier period. Future work with larger corpuses may be able to separate works by genre and thus assign dates according to a method appropriate to the genre in question.
1.4 Survey of the Field

1.4.1 CHILD MALTREATMENT IN ANCIENT ROME

Laes (2008), researches child slavery in Ancient Rome by looking at written sources. He shows that there was no direct correlation between a child’s status as either slave or free and their quality of life, a fact that has implications for whether or not slavery in-and-of itself should be considered maltreatment. Child slavery, and child labour in general, was an important contributor to the economy of the Roman Empire. Laes finds no evidence that the idea of child labour was a moral problem for the culture of the time: children’s human rights as a concept did not exist.

Harper (2013), in his book on sexual morality in antiquity, considers evidence on the sexual abuse of children. In Rome, unlike the Greek city-states, freeborn boys were strictly protected from pederastic relationships by law and Roman culture in general abhorred pederasty as traditionally practiced between freemen in the Greek cities. Slaves, however, had no such protection: life as a slave would often mean constant sexual abuse and slaves, of whatever age, were always at hand, and so were seen as a convenient avenue for sexual release. Harper’s work shows that sexual abuse of children was commonplace in ancient Rome and, regarding slaves specifically, was not regarded as morally problematic.

Regarding infanticide, Obladen’s (2015) research focuses on researching written laws to better understand the proclivity of ancient societies towards this behaviour. Law dating from the earliest years of the Roman Republic gave the father of the family the right to kill his children; these laws remained relevant for many centuries after. Deformed infants, or infants deemed abnormal or weakly, were killed, typically by drowning in a bucket. However, Obladen warns that the frequency of infanticide is hard to estimate due to the overall high infant mortality rate.

Often unwanted children were exposed instead of being killed directly. Harris (1994) considers the evidence about child-exposure in the Roman world. Importantl, he warns that such evidence cannot be "translated into numerical form". His review of the research in this area shows that exposure did not always lead to the death of the child, although this was often the case. Children who survived exposure were typically “rescued” into exploitation. Harris also points to an apparent sex-imbalance in the victims of exposure, also addressed in a number of studies referenced in Obladen (2015). There has been considerable work done the question of sex-imbalance in the context of exposure and infanticide in the field of archeology. In a review of work in this area, Gowland, Chamberlain and Redfern (2014) conclude, however, that the evidence for a sex bias is not clear, and also warn that much of the supposed archeological evidence for exposure and infanticide is hard to interpret: many buried children understood in some research to have been victims of infanticide may have in fact died due to natural causes. In an investigation
of a rare unambiguous case Faerman (1998) used DNA evidence to determine the sex of a number of corpses among a group of some 100 victims of infanticide found beneath a Roman bathhouse. The authors speculate that the high frequency of male victims was due to the bathhouse’s use as a brothel—female children could be used in the future to support the trade.

Research on the physical abuse of children in antiquity is sparse. In a chapter on Corporal Punishment in the Ancient School, Bloomer (2015) discusses the harsh corporal punishment inflicted on children in ancient Graeco-Roman society. While the norms of those societies regarding punishment would now be seen by many to be abuse, they were not considered as such by people at the time. This in turn suggests that physical abuse was likely a normal part of childhood in antiquity though, as Bloomer points out, there were cultural limits on the physical abuse of free children, at least.

Miller (2005), in a book that focuses on care for orphans, argues that orphaned and abandoned children were largely neglected, at least on an institutional level, by pre-Christian Roman society, drawing on a large body of literary evidence.

1.4.2 EARLY CHRISTIANITY AND THE CHILD

In a broad introduction to the subject of childhood and early Christianity, Leyerle (2013) concludes that while Christianity made an impact on children’s lives in general, in many ways their lives were nonetheless unaffected. In terms of work with more in-depth consideration of Christianity and the child, Strange’s (2004) book Children in the Early Church provides a non-technical overview of the cultural context surrounding childhood for the Early Church, and considers the relevant writings of the New Testament in this light, while Francis (2006) presents a collection of essays on the image of childhood presented in the New Testament and within the cultures (Jewish and Graeco-Roman) that directly influenced early Christianity. In the most comprehensive overview of the subject to date Bakke (2005) argues that the concept of childhood as it has been understood ever since originated with early Christianity. His book When Children Became People looks at the pre-Christian understanding of the child; Patristic teaching on the subject; attitudes (both pagan and Christian) towards infanticide, exposure, and sexual relations with between adults and children; children’s education; and the participation of children in religion and religious ceremony.

1.4.3 SENTIMENT ANALYSIS

Pang & Lee (2008) provide a survey of the various techniques that can be employed to automate sentiment analysis of large corpora, an area that had begun to be of particular interest at that time due to, as they argue, the rapid growth in the amount and availability
of opinion-containing text resources available online, thanks to the growth of social networking platforms such as Twitter. While fact-based analysis techniques had long existed, sentiment analysis of this kind was relatively novel at the time. Since then, virtually innumerable papers have been published focusing on the application of, technical challenges presented by, and various approaches to sentiment analysis. Many of these look at sentiment analysis as a tool within a specific domain, for example Alexander et al (2011) and Maral (2011), which focus on movie reviews; Xia et al (2014) and Duyu et al (2014), which focus on social media; and Qingxi and Ming (2014) and Myel et al (2011), which focus on customer reviews of products.

Other papers are concerned with various technical issues. In one of the earliest widely-cited papers on the subject, Wilson et al (2005) consider the challenges present by context polarity, that is how the polarity of an individual word is affected by its context, for example by negation (e.g the word “good” normally suggests positive sentiment as in the phrase this is a good movie but the polarity is reversed if the word is negated, for example in the phrase this is not a good movie; this problem is discussed in more detail below). Liu (2010) considers challenge presented by subjectivity, that is the problem of delineating between facts and opinions in automated digital analysis of texts.

A third area of research is the various approaches to sentiment analysis available to the researcher. Taboada et al (2011) present research on lexicon-based methods, whereby a dictionary of words, pre-annotated for polarity, are combined with consideration for negation and intensification to analyse the sentiment expressed in a given text, whereas Boiy (2009) looks at more sophisticated machine learning approaches, including the Support Vector Machine, Multinomial Naïve Bayes, and Maximum Entropy techniques.

1.4.4 THE APPLICATION OF SENTIMENT ANALYSIS TO NOVEL DOMAINS

Sentiment analysis and sentiment analysis like tools have been applied in novel domains, that is, outside of the typical domains of application which include movie and reviews, social medial, and customer reviews of products. Spruognoli et al (2015) present an online platform called ALCIDE which facilitates the use of sentiment analysis on historical texts, and also present the results of a first experiment with the use of the their tool, applying it to a corpus of Italian historical texts. Mohammad (2011) applies sentiment analysis and visualisations to track and quantify the presence of various emotions in individual novels and across a large corpus of both fairy tales and novels. He introduces the concept of emotion word density, and shows that fairy tales, in aggregate, feature a much wider range of emotion word densities than novels.

---

3 Twitter was founded in 2006
1.4.5 SENTIMENT ANALYSIS AND WORKS IN TRANSLATION

One of the challenges for the research presented in this paper is doing sentiment analysis like digital analysis on texts in translation rather than in their original language. Some research has been done on doing this in an automated fashion, i.e. performing machine translation of a non-English language text to first translate it into English, and then performing sentiment analysis on the translated text. Balahur and Turchi (2012) apply this technique with success, using the latest machine translation tools. Their work follows similar research by Banea et al (2008) who were interested in this approach due to the large amount of work and resources that have been dedicated to English language sentiment analysis, such as the development of lexicons and understanding of the use of negations etc. Brooke, Tofiloski and Taboada (2009) take an alternative approach, modifying the analysis tool, rather than the text, to perform cross-linguistic sentiment analysis on Spanish texts. Looking at Arabic social media posts, Salameh, M., Mohammad, S., & Kiritchenko, S. (2015) compare both these approaches, and favour the latter, developing a new tool aimed at Arabic sentiment analysis in the process.

2. Historical and Technical Theory

2.1 Defining Child Maltreatment

The Child Abuse Prevention and Treatment Act, passed in 1974 in the United States, aimed to protect children from maltreatment of two kinds: abuse and neglect. At the time, however, it was noted by both practitioners and scholars that these terms were not clearly defined (Hutchison, 1990).

As Hutchison notes, different groups use different definitions of maltreatment according to their aims. If this is true at one moment in history, then differences in definitions are likely to be even more apparent across large historical time periods. Hutchison identifies four different purposes, which each involve different ways of understanding maltreatment; two are of note here.

In the area of social policy and planning, definitions of maltreatment are part of what determines how solutions to the problem of maltreatment are developed by those making social policies. Hutchison points to the work of one scholar in this area, David Gil, who offers the definition “non-accidental physical attack or physical injury, including minimal as well as fatal injury, inflicted upon children by persons caring for them”. A second, broader definition offered by the same scholar is that “abuse of children is human-originated acts of commission or omission and human-created or tolerated conditions that inhibit or preclude unfolding and development of inherent potential of children”.

Page 10 of 58
Regarding research, Hutchinson maintains that there is a “definition dilemma” that has alternatively been solved by either adopting definitions provided by agencies researchers would consider “responsible”, or else by ad hoc specifying the behaviour or conditions that the researcher regards as properly considered maltreatment. However, the second approach can make it difficult to meaningfully compare findings from across the body of literature, as differing definitions are employed.

Hutchinson cites scholars who conclude that it is impossible to define child maltreatment satisfactorily. However, she notes there is a consensus among lay people that there are “four types of maltreatment ... serious enough to warrant coercive intervention: sexual abuse, serious physical injury, leaving a young child unattended, and fostering delinquency.”

The problem in applying such modern definitions, whether academic or lay, of maltreatment to historical research is obvious. Scholars concerned with maltreatment in the modern, Western world are unlikely to explicitly include being sold into slavery or infant exposure, just as two examples, within their definitions. While such things are not unheard of in the modern world they are both rare enough and extreme enough to not warrant explicit inclusion, despite being among the most common forms of recorded child maltreatment in antiquity.

Scannapieco (2005) contends that adopting a historical definition of maltreatment would also be insufficient, as behaviour the ancients viewed as normal and commonplace modern people would typically find abhorrent. Instead, when considering abuse, both in the historical context and in modernity, she offers four broad categories: physical and sexual abuse, neglect, and emotional abuse. Using these categories as a guideline, and with reference to the ancient literature, I identified a number of more specific types of maltreatment, each of which fits into one or more of these categories. These types of maltreatment are: abandonment; murder/infanticide; being exposed to highly traumatic scenes or situations; experiencing captivity, enslavement, or imprisonment; exposure to extreme danger, such as war or disaster; general emotional abuse; neglect; experiencing extreme poverty (which, I suggest, can be understood as a form of societal maltreatment); kidnapping; physical abuse and harm; sexual abuse; and child sacrifice or maltreatment due to involvement in a religious ceremony or rite.

2.2 The History of Child Maltreatment

2.2.1 HISTORICAL DEVELOPMENTS IN THE TREATMENT OF, AND ATTITUDES TOWARDS CHILDREN

Few scholars have attempted to build a comprehensive theory of societal attitudes towards children and childhood over the course of history. Only two such theoretical frameworks are widely cited, and both remain controversial. The first belongs to Phillips
Ariès and was presented in his work L'Enfant et la vie familiale sous l’ancien régime, known in English translation as Centuries of Childhood (Ariès, 1965). Ariès theorises there are two distinct and contrasting eras in the history of societal attitudes towards children, each largely incompatible with the another. Focusing particularly on French society, he conceives of the “traditional” era (in contrast with the “modern”) as existing from Antiquity, through the Middle Ages, and right up to the beginning of industrialisation. For Ariès, traditional society did not recognise children as being in any privileged way distinguished from adults, providing as evidence the language used to describe children and childhood, how they were depicted in art, the clothes they were given to wear, how they were expected to amuse themselves, and even attitudes surrounding children and sexuality, among others. After surviving the first seven years of their life, a period in which they were too fragile, and too likely to die, to be recognised as counting as fully human, children would participate in life much as adults did (Wilson, 1980).

Ariès finds a change in these attitudes evident in the seventeenth century, where he sees an emergence of two previously unknown societal attitudes towards children. The first, known to historians more from writings that are critical of it than those voicing support, was that of an explicit affection and enjoyment of children, now well-known in modernity—one that celebrated children’s innocence. The second was an interest in the psychological implications of a family’s approach to the moral and intellectual upbringing of a child. For Ariès, this era marks the beginning of a concern for the treatment and psychology of children that had not existed hitherto. Again, the main justification for this theory is that children, as Ariès argues, did not exist as a clearly-defined category of human beings in the popular consciousness before modernity: they were simply smaller adults (in the same way that, in modernity, we do not have a special concern for the psychology and well-being of thirty-eight-year-olds: we see no reason for them belonging to their own, privileged category). One of his main lines of evidence in support of this position is the depiction of children in the iconography of the Middle Ages: they are shown as, quite literally, smaller adults. Later considerations of Ariès’ theory find such evidence, as well as much of his overall work, unconvincing. While it may well be true that the societies of Antiquity and the Middle Ages did not categorise the stages of life in the same way as a typical modern, Western individual might, they did in fact recognise life as having distinct phases. Moreover, while concern for the plight of children may have increased in the era following the Age of Enlightenment and then during industrialisation, alongside changes in the definition and boundaries of what one might consider a “child”, it seems unlikely that childhood was simply discovered at this point in history, as Ariès suggests (Wilson, 1980).

A second and equally debated work, in part written in responses to Ariès, is Lloyd deMause’s The History of Childhood (1995), which opens with the memorable assertion that:
The history of childhood is a nightmare from which we have only recently begun to awaken. The further back in history one goes, the lower the level of child care, and the more likely children are to be killed, abandoned, beaten, terrorised, and sexually abused.

DeMause and his fellow “psychohistorians” have long held that, historically, there have been different, overlapping, eras in the treatment of children, with an overall trend of progressive improvement. According to deMause’s theory, Antiquity was typified by “infanticide” as the mode of child care. During the Middle Ages, this gave way to the “abandonment” mode, followed in turn by the “ambivalent”, “intrusive”, “socialising” and finally the “helping” modes from the mid-twentieth century onwards (Pletsch, 1975). While it may seem odd to argue that infanticide or abandonment can be interpreted as a form of child care (as deMause sees it, these modes were the normal approach to resolving child care “anxieties” in their respective eras) the work does identify changes in prevailing attitudes towards children, and in the treatment of children across history.

Both Ariès and deMause theorise significant historical development and change in the way societies have seen and treated children. While their models of this change are much debated, the fact that change has indeed occurred seems to be widely accepted (Wilson, 1980). Importantly for the present work, however, neither identify the Christianisation of Europe as a particularly important moment in this history of change: consideration of this link is found in works dealing more specifically with that question.

2.2.2 LINKS BETWEEN CHRISTIANITY AND CHANGES IN THE TREATMENT OF CHILDREN

Bakke (2005), in considering changes in attitudes and behaviour towards children in Antiquity, and the link between such changes and Christianity, looks at the early Christians’ threefold cultural heritage: Greek, Roman, and, given that Christianity was a Jewish sect, Hebrew. For both the Romans and the Greeks—all the way from Homer to Cicero—children were understood to lack logos, that is speech, rationality, or reason. They shared this defect with women, slaves, and barbarians. Given their perceived stupidity, irrationality and weakness, children were not held in high regard. The image of the child was often appealed to, in literature, as an example of what an adult male citizen should not be. Children were to be punished severely, praised only for their malleability and potential (if they were boys; of girls, little was expected) and found themselves, while they remained children, at the margins of society alongside, again, women, slaves, and the

4 Plato, for example, often included them in such a group alongside these other “unimportant” social actors.
elderly. Grief at a child’s death was to be admonished. The only justification for grieving a dead child was as a wasted investment: a child’s principal worth was as an investment for the future. Moreover, in many cases children the investment was considered not worth making. In antiquity, the majority of people were very poor and raising a child was, as it is in any age, an expensive undertaking. To deal with unwanted pregnancies, alongside highly dangerous methods of abortion, Greco-Roman societies practiced the custom of expositio (ekthesis in the Greek-speaking world), or the ‘exposure’ of children. While some exposed children certainly died many would have been found and rescued, but what ‘rescue’ typically meant in this context is also a matter of debate. No doubt some children found themselves looked after as if they were the natural children of those who found them (a trope featuring in a number of antiquity’s myths), but many foundlings would also have become slaves. As slaves, violence and threat of violence would be a normal part of their lives. Worse, from a modern perspective at least, was the all-to-common reality of sex between children—both slaves and, in the Greek world especially, free—and adults. Bakke quotes the above-mentioned Llyod de Mause, who writes: “the child in antiquity lived his earliest years in an atmosphere of sexual abuse. Growing up in Greece and Rome often included being used sexually by older men.” Sexual abuse of (particularly enslaved) children was routine in Ancient Greece and Rome, but that appears to have been all but universally accepted. For Bakke this, although only one area of children’s maltreatment, perhaps more than anything else illustrates how radically attitudes towards children changed with the onset of the Christian Era. It should be noted that there were voices of dissent, critical of society’s attitude towards sex, that came from outside the Christian community. The generally prudish Plutarch, for example, had views on sex that in some ways anticipated later Christian norms (Walcot, 1998), as did many of the Stoics. Notably, too, Philo of Alexandria was highly critical of his own city’s eroticised culture—the Jews, he claimed “know no sexual intercourse with other women but enter marriage as pure men with pure virgins”. Anything outside of this “special set of customs and norms” practiced by the Hebrews was unacceptable (Harper, 2013).

---

5 This should be understood in context, however: studies suggest around half of all children failed to reach their tenth birthday (Bakke, 2005); moreover, as Bakke points out, the fact that parents had to be “told off” for grieving their dead children demonstrates that some parents, at least, did grieve.

6 Athenagoras of Athens, among other Christians, equated exposure with murder (Bakke, Loc. 1832); I will return to Christian perspectives on exposure below.

7 De Mause, “The Evolution of Childhood” in The History of Childhood, p. 43. It should be noted that the Bakke does not present this quote uncritically: his criticism amounts to identifying the term ‘sexual abuse’ as an anachronism in this context.
Harper argues it was the writings of St. Paul, however, that turned the culture of antiquity regarding sex, and children’s welfare more generally, completely on its head. It is telling that the early Christian community coined or repurposed a number of terms to describe their views, as it suggests the vocabulary inherited from their surrounding culture was deemed inadequate to express these radical new ideas, for Harper demonstrating how much of a departure from the contemporary norms their opinions were. Among these was the compound word *paidophthoria*, meaning “the violation of children”, a word not found in sources before the emergence of the early Church. Eventually, Christianity’s stern rejection of coercive sex, including sex involving slaves, was codified not only within the Church’s own doctrine, but also in Roman law.\(^8\)

Bakke (2005) argues that Christianity, in inheriting key Jewish values regarding humankind, offered a totally new way of thinking about children. Regarding the education of children, for example, John Chrysostom, the Archbishop of Constantinople, wrote that the aim should be the revelation of God’s royal image for “Man is the image of God”. Children were no longer marginal actors, important only in terms of potential and as an investment for the future: they were the made in the image of God.\(^9\) Bakke notes that while both Gregory of Nyssa and Augustine write at length on the “suffering and death of small children” but there is

No evidence that the Greco-Roman pagan tradition ever regarded the sufferings of small children as so problematic that moral-philosophical treatise were devoted to this question. On the contrary, writers such as Cicero and Seneca criticise parents and especially fathers for grieving over the deaths of their small children.

Jesus had used small children as positive examples in the gospels. In the past, as noted above, children had typically been used as examples of what grown men should not be like. But the Church fathers recognised that Jesus held up childhood as representative of many things Christians, grown men and women among them, should imitate: simple faith, 

\(^8\) In around 380 AD, the Apostolic Constitutions would explicitly formulate what had already been taught by Christians for centuries: “Do not violate children, for contrary to nature is the evil born at Sodom, which was laid waste by the fire of God.” By this time, with the Christianisation of the empire, Christian views on sexual ethics had begun to be codified in law: 428 AD saw the Christian emperor Theodosius II ban coercion in the sex industry and by the time of Justinian (AD 527-565) it seems that “men could be accused of sexual crime even by slaves, signalling a total breakdown of the ancient sexual order” (Harper, 2013).

\(^9\) In Eastern theology in particular, that children are created in God’s image is especially emphasised. For Gregory of Nyssa the implication is that “the child shares in God’s life” (Bakke 2005).
sexual innocence, and indifference to the temptations of “mammon” (Bakke, 2005). Bakke argues it is not surprising that Jewish and Christian writers were outspokenly critical of the way their societies treated children: if children are created by God and made in His image, then clearly they should be protected.

While the case for a change in attitudes is strong, it is also necessary for the present research to consider whether these new attitudes resulted in an actual change in behaviour, resulting in less maltreatment of children. Some evidence that it might is found in contemporary writings. As an example regarding exposure, the apologetical work known as The Epistle of Mathetes to Diognetus says simply that, while Christians in general follow the customs of whatever place they find themselves in, “they do not destroy (expose) their offspring” (Chapter V). Bakke argues that in practice, too, this was the case, pointing out

\[ \text{the fact that the exposure of children does not feature in polemic within the church indicates, at the very least, that this was not very widespread practice. No source in the time before Constantine tells us that Christians practiced the exposure of children. This does not allow us to infer that it never happened, but it suggest it was a rare occurrence.} \]

After Constantine became the first Christian Emperor, these Christian values began to be worked into Roman law, which Bakke (2005), Harper (2013) Miller (2003) and Leyerle (2013) all argue resulted in more widespread practical effects. By the time of Valentinian, the law had banned exposure altogether. Roman law on exposure was now largely aligned with the long-held Christian attitude towards the practice as expressed by Basil of Caesarea, who says it should be “judged as for murder”.

The changes that came about as a result of the new Christian ethics were not only reactive—that is, designed to protect children from exposure, straight-up murder (in the case of the weak and deformed), and sexual abuse, all practices considered normal in earlier Greco-Roman culture—they also represented positive moves made to improve the lot of the least privileged children in society. Miller (2003) argues these changes anticipated the children’s welfare we have today, offering evidence that, in the West, state-led children’s welfare, only began with the Roman Empire’s conversion to Christianity in the fourth century.

There is strong, mostly written (Grubbs, 2013), historical evidence, then, that the onset of the Christian Era may have represented a watershed in attitudes and behaviour towards children, and thus a decrease in their maltreatment, in Antiquity.
2.3 Sentiment Analysis and Similar Tools

2.3.1 APPROACHES

There are two main ways to approach sentiment analysis or similar digital analysis of a text or texts. The first is a lexicon-based approach, which, as the name suggests, involves calculating the sentiment of a text based on reference to a lexicon, wherein each word has a score for polarity. The second involves building a machine learning classifier, which is trained on pre-tagged texts or sentences (Taboada, 2011).

When using a lexicon-based approach, lexicons can be created manually or by a combination of manual initial list creation and automated expansion of the manually-created list by algorithmically including synonyms and related terms. These lexicons represent a large table wherein each entry is a word with a score for polarity (either positive or negative). Then, when analysing a text, each word from the text is extracted, looked up in the lexicon, and, if found, contributes (either positively or negatively) to the overall aggregate score of the text in question. Typical applications of lexicon-based models assume that individual words have a context-free “prior polarity” that can be given numerical score. This numerical score is sometimes either +1 or -1, or otherwise exists across a spectrum from very negative (typically scored -5) to very positive (typically scored +5).

Lexicon-based approaches will also often take into account intensification, which refers to intensifiers: words that magnify or diminish the effect of words coming after them, for example “extremely” or “slightly”. Recognising the juxtaposition of intensifiers and sentiment words affects the score of the sentiment word, which in turn will affect the aggregate score of the text.

Secondly, negation is typically taken into account, as touched on above. Taboada (2011) offers some examples of negation in action:

- Nobody gives a good performance in this movie. (nobody negates good)
- Out of every one of the fourteen tracks, none of them approach being weak and are all stellar. (none negates weak)
- Just a V-5 engine, nothing spectacular. (nothing negates spectacular)

As negators often don’t occur immediately before the term they negate, any lexicon-based algorithm looking to effectively consider their impact must be more sophisticated than simply considering immediate juxtaposition. Other features of the text, such as the presence of modals and the subjectivity or objectivity of what is being expressed may also be taken into account.
The second group of approaches—those involving trained, machine learning—work by training an algorithm on pre-tagged data sets: the algorithm extracts certain information from a sentence and, during the training phrase, is then told by the programmer whether that text expresses positive or negative sentiment. By doing this many times in succession, it can “learn” what kinds of information correlate most strongly with a positive result. Once training has been completed, it can then decide to a fairly high degree of accuracy—typically higher than a lexicon-based approach—whether a text is positive or negative. However, machine learning approaches are only useful at analysing the kinds of texts that they have been trained on; if exposed to texts outside of the training domain, they can sometimes perform little better than chance (Taboada, 2011). Besides the potential for greater accuracy, the advantages of machine learning approaches include a better understanding of broader context, and the ability to offer a probabilistic score (rather than a fixed score as with lexicon-based approaches) which reflects the uncertainty inherent in any such analysis (Boiy, 2009).

The first step of building a machine learning algorithm is to have it identify features, or information, present in a text. Such features would include whole words, stems (i.e the root lemma of a word), parts-of-speech, negation, and distance between sentiment words and a target (Boiy, 2009). Various different classification algorithms can then identify how these features correlate with a given result, through training, and give a probabilistic assessment of whether a given, previously unseen, text has negative or positive sentiment, or whatever quality it is the classifier is looking to assess.

2.3.2 INTERPRETATION OF RESULTS

The algorithm will provide a numerical figure for what percentage of texts in a given decade or century feature the maltreatment of children. These figures will then be quantitively analysed to determine whether or not there is a significant change after the onset of the Christian era. To determine significance, the data will be assumed to follow a normal distribution.
### 2.4 Study Design

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Is this change reflected in the (extant) writings produced over this period, and if so in what way?</th>
<th>Are SA-like tools useful for the analysis of historical corpora, and what can they contribute?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was there a reduction or change in the maltreatment of children coincident with the beginning of the Christian Era, and in what ways did this change manifest?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Research Type**
- Quantitative, Deductive

**Data Collection**
- 6,565 sentences mentioning children, extracted from 418 public-domain texts in English translation, written in the period 88 B.C to 550 A.D, tagged by year.

**Method**
- Machine-learning naive Bayes classification of individual sentences to algorithmically determine whether or not they feature child maltreatment.
3. Material and Method

3.1 Corpus

I decided to focus on the period from the beginning of the first century B.C to the end of the sixth century A.D. This time period is fairly neatly intersected by the emperor Constantine’s conversion to Christianity, thought to be in 312, which marked the beginning of a rapid growth in both the number of people converting to Christianity, and its influence throughout the empire.

To put together a corpus representative of this period of history, I had two choices. The ideal choice would be to collect sources in the original languages, that is, Latin and Ancient Greek. Sadly, my almost non-existent reading abilities in either language would make developing an effective algorithm for the analysis of original-language texts extremely challenging (particularly in preparing training data if a machine-learning algorithm is employed), so I elected instead to look at texts in translation. This, in turn, limited my choices to translations that are in the public domain. I was able to identify and download 418 texts from the period with public-domain translations, covering the dates 88 B.C to 550 A.D.

Relevant texts were identified with reference to https://topostext.org, which includes links to the downloadable, public domain documents which I used for this research.

3.2 Identifying Relevant Sentences

I picked out relevant sentences from the corpus by writing a very small and simple program that crawled all of the texts, identifying any phrases which featured the words “child”, “children”, “infant”, “babe”, “baby”, or “foundling”, and then saved each relevant sentence into a database, tagged with the date of when the text was written. The program found 6,565 such sentences within the corpus of 418 texts.

3.4 Developing an Algorithm

3.4.1 SIMPLE TAGGING

My first experiments in analysing the data involved using the NRC Emotion Lexicon, also known as “EmoLex”, to score each word in a sentence according to its emotions. Emolex provides a list of English words, with each of them scored for eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust). The scores for each word were obtained manually, via crowdsourcing, with multiple individuals scoring each word. Each
individual was originally presented with a test word. If they “failed” in scoring this test word appropriately, any further contributions were ignored (Mohammad & Turney, 2012).

In the resulting Emolex database, common English words (of which there are 14,182 in the lexicon) are given a score of 1 or 0 for each of the emotions, depending on whether this word is associated (1) or not associated (0) with that emotion. For example, the word “abandoned” has the following associations:

<table>
<thead>
<tr>
<th>EMOTION</th>
<th>His</th>
<th>Sick</th>
<th>Child</th>
<th>Was</th>
<th>Abandoned</th>
<th>AVERAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>anger</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>anticipation</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>disgust</td>
<td>0</td>
<td>1</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>fear</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>joy</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>sadness</td>
<td>0</td>
<td>1</td>
<td>x</td>
<td>0</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>surprise</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>trust</td>
<td>0</td>
<td>0</td>
<td>x</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

I used the Emolex dataset to score each word, in each of my sentences. I then averaged the scores of each word, excluding the target word (i.e. the word associated with children, e.g. “baby”), across an entire sentence to get an overall mean score for the sentence (words not found in the lexicon are not included in the score so do not affect the average). For example, with the sentence his sick child was abandoned, the average would be calculated like so:
Unfortunately, this approach gave results little better than chance, a fact ascertained by comparing the results of the algorithm with manually-tagged sentences. While it would certainly be very interesting to look at the individual emotions, and their associations with my target words, the limited dataset of less than 7000 sentences (by digital text analysis standards, a very modest amount of data) meant that the results were mostly noise. It became clear that it was necessary to narrow the focus.

I turned next to AFINN–1651, a lexicon which considers only the negative/positive polarity of words, instead of considering their emotions. AFINN–165 is frequently used in SA research, for example in the research conducted by Howell & Melenbrink (2018). Initially my approach to analysis was similar to the one described in their paper: the AFINN–165 provides a lexicon of English words, each with a score for valence represented by an integer within the range –5 to 5; a score of five would mean the word has a very high psychological valence, for example the word “thrilled”, which indicates highly positive sentiment. The word “tortured”, in contrast, is given a score of –4. The vast majority of words in English are considered neutral, and so are not given a valence within the Lexicon. Such words are automatically scored zero. Following Howell & Melenbrink’s approach, the score for all of the tokens (words) is averaged, according to the number of tokens, for a mean valence, which itself would be in the range –5 to 5 (although, as the authors point out, in practice most sentences will have a score between positive and negative one, given the high prevalence of neutral words).

This approach also provided results that, while perhaps slightly better than the attempts made using the Emolex dataset, still suffered from a very high number of false positives. Refinement of the (as yet extremely simple) algorithm was called for.

Wilson, Wiebe and Hoffman (2005) discuss the importance of considering contextual polarity in developing successful text analysis programs. As they explain in their paper, the approach outlined above looks only at a priori, that is out-of-context polarity of a given word. For example, the word “thrilled”, as mentioned above, has a positive polarity when considered outside of any context. However, when considered in context, the meaning it contributes to the phrase or sentence within which it appears may change dramatically, for example, I was not at all thrilled to learn of their arrangements.

A simple averaging of the polarity of individual words in for this sentence (using the AFINN–165 lexicon) would give a score of +0.45, even though this sentence in fact expresses negative sentiment. In this example, this is because the word “thrilled” is negated by the word “not”, that appears three words prior. Negation of this kind may be due to immediate juxtaposition of a negative word (not thrilled), or the negation can appear somewhat removed from its subject, as in the example above. Wilson, Wiebe and Hoffman also point to a trickier occurrence where apparent negation actually serves to make the meaning of a word more intense, rather than, in fact, negate it: they give the example phrase “not only good but amazing”. The modality of a sentence can also affect
its polarity: “it should be a good movie” does not represent a ringing endorsement of its subject but an algorithm that ignored modality would likely assess it to be.

My simple tagging/averaging algorithm was improved by taking these ideas into account. Still, however, accuracy was at best no higher than 65%.

3.4.1 NAIVE BAYES

I had initially decided against taking a machine-learning approach due to the workload associated with manually tagging a training set, and the greater complexity of developing the algorithm. Machine learning requires human (in this case, me) sorting of typical input/output data from which a program can “learn” what is expected of it. In the versions of the program outlined above, the human input had come from previous, crowdsourced, teams of human taggers that had assigned a positive or negative polarity (or emotional score, in the case of Emolex) to individual words. This alone, however, had proved inadequate for the development of an sufficiently capable algorithm.

To create a training data set for this project, I randomly selected 1000 sentences from the 6,565 sentences already taken out of the larger corpus. I then read each sentence and manually tagged it with either a one or a zero, depending on whether it featured children's maltreatment or not. For example, the sentence:

Many high honours have been granted them by the commonwealth, as a result of which they feel no desire either for marriage or for children.

Was marked zero (not featuring the maltreatment of children), while the sentence

Their wives and children were carried away into slavery by the Romans and the city was plundered and burned.

Was marked one (featuring the maltreatment of children).

To process this data I wrote an implementation of a Naive Bayes machine learning algorithm. Naive Bayes is typically used in digital text analysis applications like email spam identification and often in sentiment analysis, too. Naive Bayes, as opposed to more sophisticated Bayes machine-learning algorithms, assumes each probability concerning the occurrence of a given fact about a subject is independent of the others, hence “naive” (McCallum & Nigam, 1998). For example, in a text classification program designed to identify email spam, the word “meet” would individually contribute to the probability of the text being span, independently of the word “singles”. While this assumption is often not true in real life—presumably “meet singles” is a commonly-occurring bigram—Naive
Bayes classifiers nonetheless perform very well, and have the advantage of being easy to implement.

As McCallum and Nigam explain, taking spam-identification as an example, a typical Naive Bayes algorithm in the learning phase looks at a number of spam emails and a number of non-spam (ham) emails and then determines the probability that the presence of an individual word indicates that a given message is spam. Then, in the classification stage, it looks at an unclassified email and aggregates the most interesting probabilities (i.e., words associated with lowest and highest probability of being spam) using Bayes’ formula to determine the probability that this message is spam. Messages that have a probability over a fixed amount (often 0.8) are then rejected as spam.

Simply using this approach would be possible for my purposes, but would require considerably more data than I have available. This is an issue of entropy, a factor which can seriously affect Naive Bayes performance (Rish, I., Hellerstein, J., & Thathachar, J., 2001). Put simply, the more variables you have in a given dataset, the more training has to be conducted in order for the algorithm to have a well-defined assessment of the probabilities involved in predicting a given outcome. As an extreme example, one can imagine a movie-reviewing app where users could use one of ten different emojis to express their reaction to a film they’d watched, as well as scoring the movie as simply “good” or “not good”. This is low-entropy, and a very small training dataset would be sufficient to predict the relationship between a given emoji and a “good” review, to a high degree of confidence (compare this with, for example, product reviews where sentences or any length in English are associated with zero-to-five star reviews—significantly higher entropy, which would thus dictate the use of exponentially more training data). The need for lower entropy, then, dictated that a simplification step be performed before training, and the same simplification step be performed during classification.

The simplification step involved first stripping any words from the sentence that were not either a) the target word, b) a “bad” negative verb or noun (ascertained by referencing a negative-word lexicon) c) a negation (not, no, neither etc.) word or d) a modal verb. Each of the remaining words was then replaced by a tag of its type, so the sentence:

\[
\text{While returning from Macedonia, before he could declare himself a candidate for the consulship, he died suddenly, survived by three children, an elder Octavia by Ancharia, and by Atia a younger Octavia and Augustus.}
\]

Would first be simplified to:

\[
\text{Could died children}
\]

Which would then be further simplified to:
After this, “facts” about each simplified sentence are compiled, which consider the order in which words appear in the sentence so, for the above sentence, the following facts would be “observed”:

- MODAL
- MODAL/BAD
- MODAL/TARGET
- MODAL/BAD/TARGET
- BAD
- BAD/TARGET
- TARGET

By going over the entire training dataset, probabilities that any given “fact” predicts that the sentence features children suffering were compiled. 800 sentences were used for training and 200 for testing, randomised each time. By tweaking the algorithm, I found that adding any additional words (for example, by tagging the presence of pronouns) resulted in worse performance—presumably because of increased entropy. However, individually considering the left-hand-side (i.e the words proceeding the target word) and the right-hand-side of the sentence resulted in improved performance, so “facts” where also tagged by position (e.g “LHS:MODAL/TARGET”). Testing demonstrated that the final algorithm could correctly identify sentences that featured children suffering with about 88% accuracy, a significant improvement on my earlier attempts.

I then used the algorithm to tag all 6,565 sentences to ascertain what percentage of sentences about children written in any given decade, or century, featured their maltreatment.
### KEY WORDS USED FOR DEVELOPING THE NAIVE BAYES CLASSIFIER

<table>
<thead>
<tr>
<th><strong>BAD</strong> WORDS</th>
<th>NEGATIONS</th>
<th>TARGETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>abandon</td>
<td>end</td>
<td>massacred</td>
</tr>
<tr>
<td>abandoned</td>
<td>enslave</td>
<td>misfortunes</td>
</tr>
<tr>
<td>abandons</td>
<td>enslaved</td>
<td>murder</td>
</tr>
<tr>
<td>avenge</td>
<td>enslaves</td>
<td>murdered</td>
</tr>
<tr>
<td>avenged</td>
<td>expose</td>
<td>murdering</td>
</tr>
<tr>
<td>banish</td>
<td>exposed</td>
<td>neglected</td>
</tr>
<tr>
<td>banished</td>
<td>flame</td>
<td>neglects</td>
</tr>
<tr>
<td>blood-stained</td>
<td>flames</td>
<td>orphaned</td>
</tr>
<tr>
<td>blood</td>
<td>flee</td>
<td>outrage</td>
</tr>
<tr>
<td>captivity</td>
<td>grave</td>
<td>outraged</td>
</tr>
<tr>
<td>capture</td>
<td>graves</td>
<td>outrages</td>
</tr>
<tr>
<td>captured</td>
<td>grief</td>
<td>perish</td>
</tr>
<tr>
<td>captures</td>
<td>hate</td>
<td>perished</td>
</tr>
<tr>
<td>corpse</td>
<td>hated</td>
<td>plot</td>
</tr>
<tr>
<td>corpses</td>
<td>hit</td>
<td>plotting</td>
</tr>
<tr>
<td>cries</td>
<td>hits</td>
<td>poor</td>
</tr>
<tr>
<td>cruel</td>
<td>hostage</td>
<td>prey</td>
</tr>
<tr>
<td>cruelty</td>
<td>hostaged</td>
<td>preyed</td>
</tr>
<tr>
<td>cry</td>
<td>hostages</td>
<td>prison</td>
</tr>
<tr>
<td>crying</td>
<td>imprison</td>
<td>sacrifice</td>
</tr>
<tr>
<td>danger</td>
<td>imprisoned</td>
<td>sacrificed</td>
</tr>
<tr>
<td>dead</td>
<td>kill</td>
<td>sacrificial</td>
</tr>
<tr>
<td>death</td>
<td>killed</td>
<td>scatter</td>
</tr>
<tr>
<td>destroy</td>
<td>kills</td>
<td>scattered</td>
</tr>
<tr>
<td>destroyed</td>
<td>lose</td>
<td>scream</td>
</tr>
<tr>
<td>destroys</td>
<td>lost</td>
<td>screams</td>
</tr>
<tr>
<td>destruction</td>
<td>mangle</td>
<td>seized</td>
</tr>
<tr>
<td>died</td>
<td>mangled</td>
<td>sell</td>
</tr>
<tr>
<td>dies</td>
<td>massacre</td>
<td>selling</td>
</tr>
</tbody>
</table>
4. Results and Analysis

4.1 Results

4.1.1 BY DECADE

Looking at the results by decade (figure 4.1) the first thing that is immediately apparent is the relative sparsity of written texts in the latter half of the time period; unsurprising, given the turmoil of those centuries compared to the stability enjoyed at the height of the Roman republic and during the early years of the Empire. The graph certainly gives no indication of any change starting in the early 300s, with the onset of Christianity, though there does appear to be a general trend downwards.

![Figure 4.1: Percentage by Decade](image)

4.1.2 BY CENTURY

At first glance, the overall downward trend is more starkly illustrated when grouping the data by century. Closer inspection, however, reveals that this apparent trend actually points to the 1st Century B.C as being particularly bloody in its depiction of children; the following centuries are all much the same. This is discussed further below.
4.2 Analysis

The data do not provide enough power for any strong intuition about distribution, so it is impossible to draw any hard-and-fast conclusions about significance. If, however—merely for the purposes of speculation—we assume a normal distribution of the data, only two
decades would be considered statistically significant ($p < 0.05$) outliers: the 70s and 80s B.C., while lying just outside statistical significance are the 50s and 60s of the same century.

When considering the data on a century-by-century basis, results are similar and we find that, outside of the first century B.C, the other centuries studied all appear to be much the same: there is no overall trend in one direction or another, and certainly no indication in a change in how often children are depicted as suffering maltreatment (by percentage) starting around the time of the conversion of Constantine, as my working hypothesis had suggested there might be. The indication, then, is that I should reject this working hypothesis.

4.3 Conclusions

The first of my research questions asked whether there was a reduction or change in the maltreatment of children coincident with the beginning of the Christian Era, and in what ways this change was manifest. Textual evidence and its analysis by historians of the era provides strong support for the idea that attitudes, at least, changed: exposure, infanticide, and sexual abuse of children were no longer acceptable in law and in the publicly-expressed opinion of key members of society, as they have been before, and the state recognised a need to provide for the poorest children in society for the first time (Helfer, 1999). Whether the reality changed is much harder to establish, as there is limited evidence from outside textual sources (Grubbs, 2013) and what elite thinkers, theologians, and lawmakers think best is not always a reflection of what, in fact, happens in practice. However it seems likely that changing attitudes and in particular changing laws would have had at least some effect on the actual maltreatment of children, but this is impossible to establish given the lack of firm, non-textual evidence, as already noted.

Secondly, I sort to better understand whether a change in the maltreatment of children, and attitudes towards children and childhood in general is reflected in the (extant) writings produced over this period, and if so in what way. While it is possible to look into the texts as a traditional historian would and find plenty of evidence for changing attitudes towards how society should treat children, and the most vulnerable children in particular, as discussed elsewhere in this paper, the specific research presented herein did not uncover any new evidence for such a change being manifest in the texts for the era in question. Indeed, other than the particularly violent texts of the first century B.C, the other centuries and decades considered all appear much the same: there is no significant trend in evidence. The very “worst” texts, in terms of depicting child maltreatment, are from 56 B.C, and the years 62, 70, and 80, of the same century.
Interestingly, every single one is, in fact, a speech from the famous Roman orator Cicero. The following extract is typical:

> So he murdered the woman, that he might not be cut off from his brother’s inheritance by her confinement; and he deprived his brother’s children of life before they were able to receive from nature the light which was intended for them; so as to give every one to understand that nothing could be protected against him, that nothing was too holy for him, from whose audacity even the protection of their mother’s body had been unable to preserve his own brother’s children.

Inspection of examples like this suggests that the high percentage of positive results found in Cicero’s speeches are not necessarily representative of a different attitude towards children belonging to his era, but rather are indicative of the fact that Cicero, as a statesman, orator, and lawyer, was frequently involved in representing legal cases, many of which involved children (in, for example, cases concerning inheritance, a very common legal question of the time), and sometimes crimes against children. This illustrates the problem with lumping together numerous different texts from different genres—histories, legal orations, theological texts, personal letters—into one corpus, and comparing them side by side. The frequency of depictions of child maltreatment in Cicero’s speeches does not, at least in isolation, indicate a different attitude or behaviour towards children in his era. Instead, a simpler explanation would be that given the kind of writings (or, more accurately, oratory) Cicero produced, it was only natural that such ideas should surface more often than, say, a later work on theology or Neo-Platonic philosophy: genres which would be unlikely to expend much ink on questions of the maltreatment of children. With further inspection, then, any case that there may have been an change in the depiction of child maltreatment unearthed via the digital text analysis techniques outlined in this paper becomes weaker.

The final aim of this research was to better understand whether SA-like digital text analysis tools could be useful for the analysis of historical corpora and, if so, what it is they can they contribute. The null result presented here, alongside the limitations of a relatively small corpus, translation, and mixed genres as discussed elsewhere in this paper might suggest that the answer is “no”. However, the fact that a relatively simple Naive Bayes algorithm was able to tag sentences in the corpus with a high degree of accuracy (around 88%) suggest that such algorithms may indeed prove to be a useful tool for the historian, to be used alongside more traditional approaches, both within the study of history and the social sciences. And while it may be the case that written data from antiquity is too sparse and varied (across time, place, and genre) for such techniques to be very useful in addressing that particular time period, it seems apparent to me it is almost
certain that in later eras where writing in individual genres (for example, scientific literature) is abundant enough that they can be considered in isolation, and data in general are much more plentiful, SA-like digital text analysis techniques will prove to be a valuable addition to the researcher’s arsenal, across a wide range of research topics.

5. Discussion

The vast majority of evidence on the maltreatment of children (and thus evidence for any changes in maltreatment that correlate chronologically with the onset of the Christian Era) are textual, with very limited evidence from archeology (Grubbs, 2013, Harris, 1994). Analysis of the textural sources, however, do seem to support the idea that there was a change in the maltreatment of children at this time. With the Christianisation of the Mediterranean world, ideas about what it meant to be a human began to change dramatically. In time, both law and opinion generally (at least as expressed in the extant literature) began to reflect this change, as Helfer (1999) concludes.

Attitudes towards how society should look after neglected children shifted radically. In previous eras, the existing cultures of the ancient Mediterranean world had offered little support for orphans. Such was the indifference shown towards neglected children that Clement of Alexandria (c. 150 – c. 215) could complain that many of his contemporaries were more concerned with exotic pets, so that “though maintaining parrots and curlews, they do not receive the orphan child; but they expose children that are born at home, and take up the young of birds, and prefer irrational to rational creatures.” In contrast with these earlier cultures, the early Christians had a deep respect for human life. Athenagoras of Athens wrote that Christians avoided watching all gladiatorial combat and executions, and never practiced abortion or exposure (Miller, 2003). And, from from the very first years of the Church’s existence, Christians were particularly concerned with providing care for orphans. This concern should not be surprising. The Old Testament has 13 references to orphans, each either encouraging care for orphans, or condemning those who fail to care for them. In the New Testament, the Epistle of James says that “Pure

10 Clement, Paedagogus 3.4.30

11 As an example, the Shepherd of Hermas, dated to the later first century or early second century—that is, very early on in the life of the Church—mentions community welfare organised by the Christians in Rome to provide for widows and orphans (Miller, 2003, p. 33). Miller also points out that Lucian of Samosata, a sceptic, satirist, and a non-Christian, could not help but notice the connection between the Christians (who he says followed a “queer creed”) and their concern for orphans and widows, again at a very early date.


Page 31 of 58
religion (thrēskeia) and undefiled before God and the Father is this, To visit the fatherless (orphanós) and widows in their affliction, and to keep himself unspotted from the world,” while Jesus reminds his disciples that “inasmuch as you have done it unto the least of these My brethren, you have done it to Me”. For Christians, to help widows and orphans was quite literally to be serving Christ. Large-scale welfare for neglected or underprivileged children, then, only arrived in the Mediterranean world with the Christianisation of the empire, beginning under Constantine. As early as 315, Constantine ordered the praetorian prefect to make payments from his own private estate, as well as the regular treasury, to parents who were too more to raise their children throughout Italy (Miller, 2003). A law of 322 extended similar welfare to a much broader part of the empire.

Regarding exposure and infanticide, a law of 322, less than a decade after Constantine had declared himself a Christian, says “if any parent should report that he has offspring which on account of poverty he is not able to rear, there shall be no delay in issuing food and clothing, since the rearing of a newborn infant will not allow any delay”. The aim was clearly to avoid the exposure of children and Bakke (2005) writes that, on this subject, “it is likely that Constantine was influenced by the Christian opposition” to the practice. By 541, the emperor Justinian had to use his Novellae Constitutiones (his new laws or, simply, “novels”) to respond to the great cultural shifts in Roman society by explicitly condemning the abandonment of children in churches.

Where abuse is concerned, laws passed by Christian emperors in the period after Constantine served to protect children from what we would today call sexual abuse, something that had been commonplace before.

Despite the results presented in this study, then, the case made by historians remains strong that the plight of children was likely improved by the Christianisation of the empire, whether regarding infanticide, exposure, slavery, or sexual abuse.

To present this picture, I’ve relied on the work of modern historians relying on extant historical texts. The emphasis within these works on Christianity could be taken as evidence of bias. However, neither I nor the authors in question have sort to argue—even

13 James 1:27

14 Matthew 25:40

15 Where previously, Romans had typically exposed children on a roadside or some other place, it seems that by the sixth century the common practice was to abandon them in Churches—reflecting how care for orphans was now seen to be the Church’s responsibility. Justinian’s novels are, in fact, an important source for understanding the scope of the philanthropic institutions that had been established by the early part of the sixth century. Alongside systems for the care and education of orphans that had already been weaned, Justinian’s writings also indicate the presence of a large network of “infant asylums” where wet nurses were paid and organised to nurse the youngest orphans and foundlings (Miller, 2003)
if making such an argument were possible—that Christianity was a necessary condition for the emergence of a culture that cared for the welfare of its children, nor to claim that concern for children is exclusively a Christian phenomenon. The experience of regions outside of the West is not the focus of the current study, so the contribution of, say, Hinduism to changes in the treatment of children across the Indian Subcontinent or the impact of the thoughts of Confucius on society’s attitudes towards Children in China are beyond the scope of what is considered here. Nor, as noted, does it seem to me that it would it be possible to explore whether Christianity was necessary for the emergence of these ideas in the West without delving into counterfactualism and indeed simple speculation, so such considerations are left to one side herein. My aim, and as far as I can tell the aim of the authors I rely on, is only to show that, in Western countries, views held to this day on the importance of society’s role in ensuring the wellbeing of children are inseparably connected with the West’s Christian past.

Regarding what, if anything, SA-like digital text analysis of the texts of the period can contribute to our understanding of this picture, however, there are obviously a number of other fairly important limitations of this study. Firstly, if a clear increase or decrease in the number of tagged sentences had been observed, what would that, in fact, indicate? A society that is more concerned by the plight of children might be more likely to write about their maltreatment; conversely, a society in which children suffer less might present the chroniclers and commentators of that society with fewer opportunities to talk about maltreatment. Another problem with looking back at this period is that the documents we still have access to represent, at least in part, the prejudice and beliefs of the society that came later: particularly post-Constantine, many writings that did not agree with the official position of the Church were destroyed; we have also lost countless texts from before that period, often for unknown reasons. Therefore, the corpus of extant documents can not necessarily be taken to accurately represent what people thought at the time but is instead, at least in part, representative of what elites in the following decades and centuries wanted people in general to be able read and have access to. The study is also limited by the fact that it looked at texts in translation. This provides a further, and perhaps more serious “filter” on what can be studied. It also, naturally, takes analysis a step further away from the original language, and accordingly all the subtlety, cultural cues etc. that may be in that language, but are lost in translation. Lastly, it considerably restricts the size of the dataset, which in turn presents its own issues for digital text analysis, as noted above.

Another problem with the application of SA-like digital text analysis (and, more broadly, any sort of historical investigation of this kind), if any changes are noted, is how to understand them in the context of the overall history of the period which is being studied. Taking our own period as an example, if a statistically significant downward trend in depictions of child maltreatment had indeed been observed, would that indicate in any
way that Christianity was the responsible force? Of course, it would not. Such an argument could be made, but evidence of this kind could only provide a small piece in a much bigger and ultimately, however well-defended, non-definitive position on the subject. It is clear to anyone with a passing familiarity with this period in history that the centuries after the Edict of Milan were a time of massive upheaval for the world of late antiquity: it was less than two hundred years later that the Western Roman Empire fell. Societal changes of any sort that occurred over this time period could justifiably be attributed to this historical earthquake, and not Christianisation, the shockwaves of which were felt for a millennia afterwards, and possibly longer. This problem, however, is well understood in the study of history, and does not represent a major challenge for the use of SA-like tools, if they are understood within their proper context and used as part of a larger argument, which draws on a number lines of evidence, as well as sound theoretical reasoning.

Despite these issues, the achievement of an accuracy of above eighty percent, even with a relatively unsophisticated algorithm, trained on a small dataset, does suggest that tools like the one presented in this paper could be useful to both the historian and the social scientist, as well as in other disciplines. While the ability of an algorithm to “understand” a text cannot (at least for the moment, and arguably ever) compete with a human being, computers can process text exponentially faster. In a matter of minutes an algorithm can read and tag more information than a researcher could read in a lifetime. This suggests that further research could be conducted to better understand how such tools might be used in the future not to replace, but to be used alongside, more traditional research techniques. Digital text analysis tool could perhaps best be used on large, more focused corpora—that is, more strictly within a certain genre and possibly considering a smaller time period—to track changes and also to identify areas worthy of further investigation in a way that it would be impossible for an individual researcher, or indeed a team of researchers, to perform unaided.
References


