

# Multimodal Metaphor Construction and Cognitive Analysis in Educational Cartoons

Genshan Wang

Chongqing Normal University, Chongqing, China

**Abstract**—Since Lakoff and Johnson proposed their far-reaching theory of conceptual metaphor, people's interpretation of metaphor has changed from a rhetorical approach to a way of thinking. In recent years, multimodal metaphors have gradually received attention from scholars at home and abroad, and have been applied to different forms of media research such as advertisements, comics, gestures, and films, thus greatly contributing to the advancement of multimodal metaphor theory. Multimodal metaphors provide a new perspective for the construction and understanding of metaphorical meaning. This study draws on 100 educational cartoons published by China News Cartoon Network from 2020-2021. Firstly, multimodal metaphors in the 100 educational cartoons published on China News Cartoon Network are classified into four categories according to Forceville's cognitive-linguistic theoretical framework, and the integration network in the cartoons is classified into four categories from the perspective of conceptual integration. This study analyses multimodal metaphors from both social and cognitive perspectives and explores the cognitive basis of comic metaphors from both formal and meaningful perspectives.

**Index Terms**—educational cartoons, multimodal metaphors, cognitive analysis

## I. INTRODUCTION

According to Lakoff and Johnson (1980), “metaphor is an important way in which humans perceive the world and also is a feature of all human thinking, present in our everyday life, language and thinking”(p.8-11). The formation of metaphors involves projections between the source and destination domains. In the process of understanding metaphors, the structure of the source domain is projected massively and systematically into the target domain, which eventually becomes part of the structure of the target domain. The meaning and structure of the source domain, therefore, determine the meaning constructed by the metaphor.

The media and channels that humans rely on for communication are called 'modality', such as language, sound, colour, image, gesture, and other symbolic systems (Zhu Yongsheng, 2007). Multimodal metaphors refer to metaphors in which the source and target domains are represented by two or more modalities, such as words and images, which provide a more accurate tool for constructing the meaning of multimodal discourse.

Cartoons are a form of graphic representation that expresses the creator's viewpoint on an event. They are usually used as illustrations in newspapers, magazines and online news editorials, and are generally easy to understand. This paper will analyze the critical metaphors of multimodal metaphorical discourse from the perspectives of cognitive linguistics and social semiotics respectively. The formal features of multimodal discourse are first described, followed by an analysis of how the meaning of metaphors is constructed using the theory of conceptual integration, and finally, the causes and effects of metaphors are explained in the light of their socio-cultural context.

## II. THEORETICAL BASIS

### A. Multimodal Metaphors

A systematic and complete theory of cognitive metaphor cannot be devoid of non-linguistic modal metaphors, as Charles Forceville (1996, 2006, 2009) pioneered the study of multimodal metaphors and has since pushed the metaphor research, which had been limited to the purely linguistic domain, to a multimodal platform. Subsequently, he further categorized the types of multimodal metaphor construction into five types: contextual metaphor, mixed metaphor, image explicit metaphor, integrated metaphor, and text-image metaphor. He and Eggertsson (2009) further proposed a broad definition of multimodal metaphor, i.e. "a metaphor that is constructed by the joint participation of more than two modalities. In this paper, we adopt the broad definition of multimodal metaphors in defining and interpreting multimodal metaphors.

EIRefaie (2009) used an empirical study to investigate the impact of readers' cultural background and individual factors on the interpretation of political cartoon metaphors of 25 culturally diverse young people who interpreted two cartoons from the 2004 US presidential campaign. Other scholars have also done extensive research on spoken language, gestures, music, and film.

Domestic studies on multimodal metaphors include studies on multimodal types and their representations (Feng Dezheng 2011), and studies on multimodal metaphors in different language categories, of which comics account for a

large proportion (Bounegru & Forceville 2011; Yu Yanming 2013; Zhao Xiufeng; Pan Yanyan & Zheng Zhiheng 2017).

### B. Conceptual Integration Theories

According to Fauconnier (1996), 'mental space is a small package of concepts that people construct for partial understanding and action when they think and talk' (p.113-129). The theory of mental space was developed to investigate the phenomenon of indirect and implicit referents in language, and it is highly effective in explaining many other complex semantic and pragmatic phenomena. For example, in terms of mental theory, the 'computer virus' involves two psychological spaces, namely the space of 'computers and programs' and the space of 'biology and medicine', although the two belong to different psychological spaces. Although they belong to different psychological spaces, the two can be mapped to create a new fusion through some elements. Later on, Fauconnier et al. extended and developed the theory of mental space by exploring a common cognitive process -Conceptual Integration. Conceptual blending refers to the process of creating a third space, the blend, based on the two input mental spaces and extracting the partial structure from the two input spaces to form the emergent structure. He elaborates on the emergent structure and argues that it is not primarily manifested in the mental space of fusion, but in the entire network of conceptual integration. In short, conceptual integration matches two input mental spaces through cross-space mapping, and selectively projects the two input spaces onto a third space to obtain a dynamic composite space.

## III. RESEARCH DESIGN

### A. Research Questions

The following specific research questions are proposed based on the above literature review,

- (1) Formally, according to Forceville's theoretical framework. How many types of construction of multimodal metaphors in educational cartoons are there? What is the frequency?
- (2) In terms of meaning. What is the distribution of the 100 cartoons according to the four types of Fauconnier and Turner's conceptual integration network? What is the cognitive rationale for this?
- (3) How can the meanings expressed by the cartoons be evaluated?

### B. Research Corpora and Methodology

A combination of qualitative and quantitative research methods was used in this study, and the cartoons were obtained from China News Cartoons. One hundred educational cartoons containing multimodal metaphors were selected from the cartoon-education category and numbered according to 1-100. The number and proportion of each metaphor type and network type were counted based on the classification of multimodal metaphors proposed by Forceville (2008) and Fauconnier's (1994) classification criteria for conceptual integration networks, respectively. We then used a descriptive approach to analyze each metaphor type formally on a case-by-case basis.

## IV. RESEARCH FINDINGS AND DISCUSSION

### A. Types of Multimodal Metaphor Construction in Educational Comics

TABLE I  
TYPES AND NUMBER OF MULTIMODAL METAPHORS

No.	Type of metaphor construction	Quantity	Proportion
(1)	Verbal-pictorial metaphors	64	64%
(2)	Target domain-implicit metaphor	9	9%
(3)	Contextual metaphor	16	16%
(4)	Hybrid metaphor	11	11%

To get a clear picture of the types of metaphor construction of educational comics, we have selected cases for discussion according to categories.

#### 1. Contextual Metaphor

The contextual metaphor is so-called because the context builds a mapping from the source domain to the target domain in this type of metaphor, and once the context is removed, the metaphor ceases to exist (Forceville 1996, 2008). on the other side is the promulgation of the national education document "Double Reduction", which conveys the metaphor that "the education sector is trying to reduce the burden of learning and increase the physical exercise of students", providing contextual information through the text labelled "double reduction" and "physical exercise, outdoor sports". Without the textual labelling, we would only see what appears to be a student playing volleyball in the playground and would have no time to interpret the direction of the policy adjustment. Then the metaphor disappears.



Figure 1: Double Increase and a Double Decrease

## 2. Hybrid Metaphor

Hybrid metaphor is an unconventional visual fusion phenomenon. Two originally unrelated source domains are mapped to the target domain by fusion. If the image context is hidden, the reader can still recognize the two objects. As shown in Figure 2, the trees and the educators in the cartoon, which are supposed to belong to very different domains and space, are morphologically fused, creating the metaphor "some schools use students as profit-making tools to extract money". The source domain is the 'money tree' and the target domain is the 'squeezed students'.



Figure 2: Money Tree



Figure 3: Long Overdue

## 3. Verbal-pictorial Metaphors

Verbal-pictorial metaphors are a phenomenon in which the source or target domain acts as a text. Thus, the text is partially or directly involved in the construction of the metaphor, and if the textual information is left, one of the domains of the metaphor disappears and the metaphor ceases to exist. This is the fundamental difference between this metaphor and others (Forceville, 1996). In Figure 3, the arched doorway represents xx University, which is presenting a termination notice to a wolf-like creature in human garb, with the word 'termination' written in bold letters on the paper. The man with the name "wolf" on his tail and a nameplate indicating "professor" on his chest represent the misbehaving teacher with a sad expression on his face. The image conveys the metaphor that "a university maintains a zero-tolerance attitude towards misbehaving professors and resolutely dismisses academics who have misbehaved. To purge the teaching community of the evil spirit". Another example is Figure 4, where the major social actors appear as umbrellas, distinguished by different colours, on top of the heads of children, with the words "Internet, family, government, school, justice, society", conveying the metaphor of "all aspects of society working together to care for the healthy growth of students". The source domain is the umbrella of different colours, represented by the textual modality of the image, and the target domain is the protection of students' growth, represented by the modality of the image.



Figure 4: The Six Protections



Figure 5: Difficult to Break Free

#### 4. Target Domain Implicit Metaphor

The educational comics explored in this study are usually illustrated from articles such as news commentaries and coupled with the fact that the exploration of educational reform is one of the hottest topics in China today, the creators often default to the fact that the reader already has some background knowledge of educational hotspots. Therefore, the contextual information about the education hotspot phenomenon does not appear in the comics, i.e. when looking at such comics alone, one usually cannot find any modal information directly characterising the target domain, but one can infer the interpretation of the specific education hotspot phenomenon due to people's social experience and awareness. For example, Figure 5 simply shows an image of a mobile phone and a wireless network signal, with a teenager, stuck on the sandwich of the wireless network signal. There are no modal expressions or education-related meanings in this figure, meaning that the target domain is not clear enough, and there is a lack of clear contextual or textual information to support the metaphor that this is a contemporary Chinese youth addiction to mobile phones and the internet.

#### B. Types of Conceptual Integration Networks in Educational Comics

Fauconnier and Turner (1994) classified conceptual integration networks into four types:

TABLE 2  
DISTRIBUTION OF CONCEPTUAL INTEGRATION NETWORK TYPES

No.	Type of conceptual integration	Number	Proportion
(1)	Simple networks	Number	16%
(2)	Mirror networks	Number	14%
(3)	Single-scope networks	Number	56%
(4)	Double-scope networks	Number	12%

##### 1. Simple Networks

Simple networks involve two input spaces, one containing an abstract organising frame containing roles, and the other containing values, i.e. the components used to fill the roles. The synthetic space integrates frame and value most simply. As is shown in Figure 9, the top right of the image is marked with the words 'After the reduction', indicating that the theme of the comic is 'education reduction'. There are two hands cutting hair directly above the subject line, with different names for different types of hair. Common sense tells us that the hair is not specifically named and that this is a reference to the specific content of education reduction. The hair of the newborn students can then grow into sports activities such as football and badminton, metaphorically reducing the burden of study and increasing the physical activities of students to promote their all-round development. By putting the policy of double reduction in education and hair together, the creator expresses the meaning that "the effect of the policy of double reduction is as immediate as trimming hair".



Figure 6: After the Reduction

##### 2. Mirror Networks

The mental space of mirror networks, i.e. the input space, the genus space and the synthesis space, has the same

organisational framework and there is no conflict between the frameworks. Take Figure 7 as an example.



Figure 7: Scramble

Figure 7 consists of two parts, the left and right parts depicting similar scenes, i.e. different doors holding the child's hand. According to the text "The library" and "Come to me, I will make you learn". The reader will know that the left part of the comic depicts a library that helps students to grow in knowledge. The right part of the comic is the same scene, but with new elements, including the colouring of the doors and the new text "Playroom" "Come to me. I'll make you have fun". The left and right sections come from an input space. The first input space contains elements such as "library, increase knowledge", while the second input space contains elements such as "game room" and "have fun". The organisational framework for both input spaces is the door. The organisational frame provides the topology for the organisation of the space, that is. It provides a set of organisational relationships for the elements in the space. The study framework organises the elements in the two spaces in an orderly manner, constructing two identical scenes. The consistency of the organisational structure makes it easier to correspond between the elements; the green door corresponds to the library, growing the knowledge of the students, etc. The elements in the two input spaces are selectively projected onto the synthetic space, constructing a mirror network with the same organisational structure in all four spaces (see Figure 8). As there is no conflict between the organisational frameworks of the four-space, the reader is easily able to find the correct relationship to carry out the compression. By mapping across space, the brown door corresponds to the playroom and to have fun.

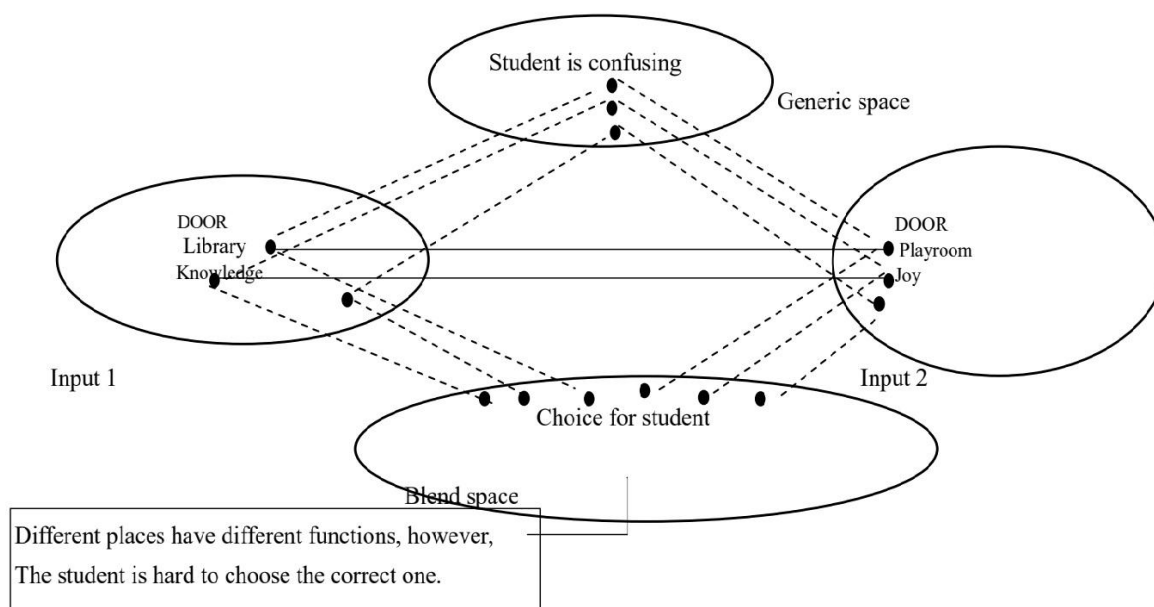


Figure 8 Schematic Diagram of the Mirror Network

### 3. Single-scope Networks

The two input space of a single-scope network has different organisational frames, and the frame of one of the input space is extended to form the organisational frame of the synthetic space.





Figure 9: Increasing Thresholds

Cartoon 12 depicts a recruitment scene in which the recruiting parties are a recent graduate and a school recruiter. The recruiter holds up a signboard to indicate his intentions. In the upper right-hand corner of the board is a sign that reads "Doctoral degree required". On the other side of the board is a primary school student studying. The cartoon thus represents the recruitment scene and a real student studying at school. The cartoon involves two input spaces. The recruitment dialogue constructs input space I1, which contains elements such as recent graduates, school staff and display boards, while the student learning scene constructs input space I2, which contains elements such as students and desks. The two input space has different organisational frameworks, with input space I1 providing the organisational framework for the synthetic space.

#### 4. Double-scope Networks

The two input spaces of a double-scope network have different or even conflicting organisational frameworks. Both organisational frameworks have an impact on the framework of the synthesis space. This is shown in Figure 10.



Figure 10: Strict Entry and Exit

The two input space of a double-scope network has different or even conflicting organisational frameworks. Both organisational frameworks have an impact on the framework of the synthesis space. This is shown in Figure 10.

This cartoon involves two input spaces. The two people at the entrance and exit of the qualification building in the cartoon constitute the first input space (input space I1), i.e. the qualification system can be a circular system that can be entered and exited, including elements such as the person carrying a school bag, the person with a diploma certificate, and two passage indications that indicate the direction. The second layer of people struggling to climb the stairs towards the dissertation, practice and defence floors constructs a second input space (input space I2), which mainly includes elements such as running people, floors with different signs and stairs. Input space I2 provides the background framework knowledge for input space I1, and the two input spaces have different organisational frameworks. In the integration process, input space I2 projects elements such as thesis, practice and defence in the process of obtaining a degree into the image space, and maps them with each other with people containing exit signs and holding diploma certificates, etc. Elements from the two input spaces are selectively projected into the synthesis space, where there are elements such as people with backpacks corresponding to exits and bachelor's hats corresponding to universities. The synthetic space then unifies the two input spaces under the constraints of the genus space, i.e. students need to pass the dissertation, practice and defence before they can receive their diploma in the process of entering the university and obtaining the qualification. It can be seen that in the dual-domain network, the organisation of the synthetic space is integrated by the parts of the two input spaces.

#### C. Metaphorical Architecture and Cognitive Analysis in the Educational Cartoons

The 100 educational cartoons were sorted and the metaphorical architecture behind the cartoons was analysed according to the representation of the source and target domains (Liang, Jing-Yu and Wang, Shao-Hua 2018), and the common architectural metaphors in the educational cartoons were put in Table 3.

TABLE 3  
COMMON ARCHITECTURAL METAPHORS IN EDUCATIONAL CARTOONS

Architecture metaphor	Source domain conceptual representation
Double increase and double decrease policy	Paper, basketball, badminton, playground
Students addicted to the internet Internet	Internet, wifi signal, mobile phones
Increasing academic requirements for teachers	Increasing academic requirements for teachers
The community protects students' development	Umbrella, hand, shield

What follows is a metaphorical cognitive analysis of the educational comic from two perspectives: the representation of the source domain and the type of metaphorical construction. Firstly, the representation of the source domain can reflect the position and attitude of the author of the educational comic. Take the metaphor of "double increase and double decrease" as an example, the increase and decrease show the two-way adjustment in the field of education. For example, the double reduction policy means that the state has adopted official documents to adjust the excessive burden of homework and out-of-school training for students in compulsory education, while the double increase refers to increasing the opportunities for students to participate in physical education, art and sports on the one hand, and increasing the time for students to study subjects such as physical education, music and art on the other. It implies that according to the different stages of national development and the new development contradictions presented in the field of education, the education sector is guiding the direction of talent training and students' personal development, changing the phenomenon of excessive academic burden and promoting the all-round development of students' moral, intellectual, physical, aesthetic and social development. The educational documents, playgrounds and students' smiles presented in the cartoon are all positive concepts, indicating that the Double Increase and Double Decrease policy has been unanimously appreciated and praised by Chinese society. The metaphor of 'students addicted to the internet, the internet, wireless internet signals and mobile phones all suggest the negative impact of the internet and smart electronic devices on students, reflecting the serious situation of the internet as a serious distraction from students' studies. These correspond to the dizzying addiction of the students in the comics and the compulsion of the scandalous electronic devices. For the metaphor of 'increasingly high educational requirements for teachers', the portrayal of school recruiters can also reflect the social reality of the education system. In some of the cartoons, school recruiters even become interview barriers themselves, and their facial expressions are exaggerated, with arrogant facial expressions, suggesting an image of arrogance, suggesting an over-saturation of demand for talent in education and competition between candidates. Facial expressions themselves are exaggerated, suggesting an image of arrogance, suggesting an over-saturated demand for talent in the education sector and fierce competition between candidates. In addition, the metaphor of 'the community protecting the growth of students' is characterised by a high frequency of images such as the umbrella, the lifting hand and the shield, which can be interpreted intuitively as positive representations; i.e. the community pays close attention to the growth and development of students and works together to resist negative factors harmful to their development. The community is concerned with the growth and development of students and is working together to counteract factors that are harmful to their development.

## V. CONCLUSION

In this study, the 100 educational cartoons are grouped into four metaphor types and four conceptual integration network types, based on Forceville and Gilles Fauconnier's theoretical framework. Respectively, with text-image metaphors and single-domain integration networks accounting for the largest share. This is related to the fact that 'uni domain networks are a prototype of the traditional source-target metaphor'. At the same time, this paper interprets the meaning and cognitive processes embodied in the cartoons in terms of both the form and the meaning of multimodal metaphors and evaluates the viewpoints embodied in the cartoons from the perspective of social meaning about the realities reflected in the cartoons. The analysis of the material shows that in the construction of the meaning of multimodal metaphors, the interpretation of meaning is not a one-way operation from the source domain to the target domain, but a multi-directional and dynamic process of construction. The distinctive feature of this dynamic construction is that it not only constructs synthetic spaces containing multiple layers of structure but also maintains conceptual connections to the network as a whole. At the same time, conceptual integration theory guides us to interpret the dynamic construction of the meaning of multimodal metaphors from a cognitive perspective better, and the combined use of the two provides a new perspective on the interpretation of multimodal metaphors in terms of form and meaning.

## REFERENCES

- [1] ElRefaie. (2009). Metaphors in political cartoons: Exploring audience responses//Forceville C, Urios-Aparisi E. *Multimodal Metaphor*. Berlin / New York: Mouton de Gruyter: 173-196.
- [2] Eggertsson G & Forceville C. (2009). Multimodal expressions of the HUMAN VICTIM IS ANIMAL metaphor in horror films. In Forceville C & Urios-Aparisi E( Eds.) . *Multimodal Metaphor*. Berlin. Mouton De Gruyter: 429-449.
- [3] Forceville C. (2002). The identification of target and source in pictorial metaphor. *Journal of Pragmatics*, (34): 1-34.
- [4] Fauconnier G, M Turner. (1996). Blending as a central process of grammar//Adele Goldberg. *Conceptual Structure. Discourse and Language*. Stanford: CSLI Publications: 113-129.
- [5] Feng Dezheng, Zhang Delu, Kay O'Halloran. (2014). Advances and frontiers in multimodal discourse analysis. *Contemporary Linguistics*, 16(01):88-99+126.
- [6] Lakoff. G, Johnson. M. (1980). *Metaphors We Live by*. Chicago: University of Chicago Press: 8-11.
- [7] Pan Yanyan,Zheng Zhiheng. (2017). A multimodal cognitive-critical perspective of defense discourse: a comparative analysis of Chinese and American military recruitment propaganda film as an example. *Foreign Language Research*, 34(06):11-18.
- [8] Yu Yanming. (2013). A study of multimodal metaphorical representations in news cartoons - types, characteristics and justifications of modal configurations. *Foreign Language Research*, (01)1-9+112.
- [9] Zhao Xiufeng, Dai Xinyang. (2016). Metaphorical scenes in political cartoons - a critical multimodal metaphor analysis. *Journal of University of Science and Technology Beijing (Social Science Edition)*, 32(05):8-13+60.
- [10] Zhu Yongsheng. (2007). Theoretical foundations and research methods of multimodal discourse analysis. *Journal of Foreign Languages* (05), 82-86.

**Genshan Wang** was born in Chongqing, China in 1996, is a Master's student currently studying at Chongqing Normal University, majoring in Subject Teaching (English).

He is currently a second-year MA student in English with a research interest in English Education.

Wang is a member of the Student Council of the School of Foreign Languages at Chongqing Normal University.