Automatic Analysis of Argumentative Discourse and Argumentation-based Human Persuasion

Dr. Stella Heras (PhD student Ramón Ruíz-Dolz & Prof. Ana García-Fornés & Dr. Joaquin Taverner)

Valencian Research institute for Artificial INtelligence (VRAIN)
Universitat Politècnica de València

Artificial Intelligence in Research and Applications Seminar (AIRA)
26th January 2023

VRAIN

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

VRAIN - Valencian Research Institute for Artificial Intelligence

Centre of excellence in Artificial intelligence research:
`https://vrain.upv.es`

VRAIN
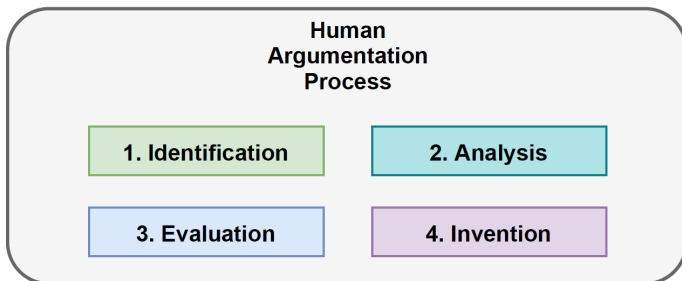Valencian Research Institute
for Artificial Intelligence

- More than 120 members (76 Doctors and 16 Full Professors)

- **TAILOR** project, key initiative for the future of AI in Europe.

- 30 years experience in **AI** and **Software Engineering**.

- Members of **Claire**, DIHs, wide experience in European projects.

- AI-focused research groups: Language Engineering and Pattern Recognition, ML, MAS, KRR, Logic Programming, Trustworthy AI, etc.

Outline

Outline

Human Argumentation vs. Computational Argumentation

- **Identification** → Look for the main conclusions and the premises that lead to them.
- **Analysis** → Detect argumentative patterns and structures.
- **Evaluation** → Measure the validity/strength of arguments based on different factors (e.g., values, topic, logic, coherence, etc.).
- **Invention** → Create new arguments and structure them in order to support and/or prove some specific idea.

Human Argumentation vs. Computational Argumentation

1. **Identification, 2. Analysis, 3. Evaluation, 4. Invention**
   - **Arg. Mining** $\rightarrow$ Identification of natural language argumentative components and their relations
   - **Arg. KRR** $\rightarrow$ Computational representations of arguments and arg.-based computational reasoning.
   - **Arg. HCI** $\rightarrow$ Generation of new arguments and adapt them for human interaction/persuasion.

Computational Argumentation Process

Argument Mining (1,2)

Argument-based knowledge representation and reasoning (2,3)

Argument-based human computer interaction (4)

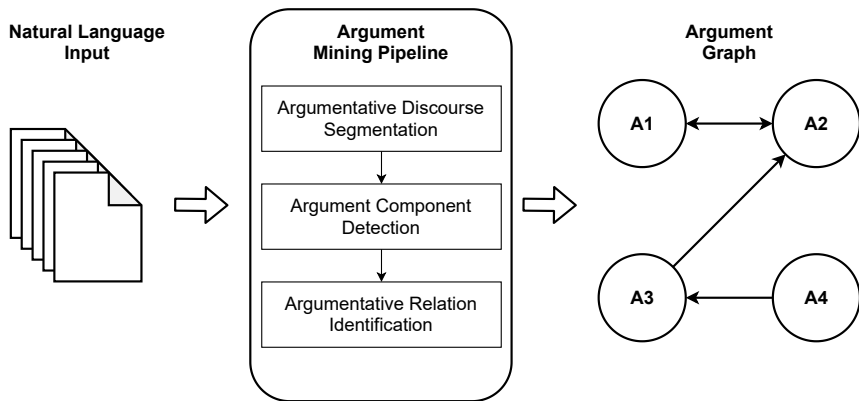# Human Argumentation vs. Computational Argumentation

Computational argumentation is a **multidisciplinary** area research:

$\rightarrow$ Natural Language Processing & Computational Linguistics

$\rightarrow$ Argumentation Theory & Formal Logic

$\rightarrow$ User Modelling & Persuasive Technology

Needs of **transversal** research that can benefit from **synergies** between different areas of research.

Outline

Transformer-based models for automatic identification of argument relations: A cross-domain evaluation[1].

We decided to use the pre-trained models that performed the best in other NLP tasks such as Natural Language Understanding, Question Answering or Text Generation:
$\rightarrow$ BERT, XLNet, RoBERTa, DistilBERT, ALBERT
Transformer encoder + Linear Layer & Softmax func.

Identify argumentative relations in natural language text:

- Default Inference (RA)
- Default Rephrase (MA)
- Default Conflict (CA)
- No-relation (NO)

Evaluate model performance in 5 different domains.

---

[1]Ramon Ruiz-Dolz et al. "Transformer-based models for automatic identification of argument relations: A cross-domain evaluation". In: *IEEE Intelligent Systems* (2021).

Political argumentation corpus for argumentative relation identification: *US2016* & *Moral Maze BBC Show* (argument maps)

	US2016	MM2012	B	E	M	P	W
RA	2744	833	128	121	205	192	187
CA	888	200	26	36	30	45	63
MA	705	156	3	25	48	41	39
NO	8055	2209	292	339	526	517	537
Total	12392	3398	449	521	810	795	826

Bank (B), Empire (E), Money (M), Problem (P), Welfare (W)

$\rightarrow$ Argumentative Relation Identification

Table 1: Performance of the models in the automatic identification of argument relations, given in macro F1-scores.

Experiment	US2016-test	MM2012	Bank	Empire	Money	Problem	Welfare
LSTM (baseline)	.26	.24	.25	.22	.24	.25	.23
BERT-base-cased	.62	.53	.40	.45	.54	.47	.53
BERT-base-uncased	.65	.56	.42	.48	.54	.50	.54
BERT-large-cased	.61	.55	.45	.49	.53	.47	.51
BERT-large-uncased	.66	.57	.47	.49	.56	.49	.57
XLNet-base	.65	.56	.44	.49	.51	.54	.55
XLNet-large	.69	.57	.44	.51	.53	.53	.54
RoBERTa-base	.68	.58	.51	.52	.54	.52	.58
RoBERTa-large	**.70**	**.61**	**.53**	.53	**.59**	**.56**	**.59**
DistilBERT	.55	.42	.33	.39	.40	.43	.39
ALBERT-base-v2	.60	.54	.49	.45	.53	.47	.51
ALBERT-xxlarge-v2	.67	.59	.50	**.54**	.56	.48	**.59**

Code: https://github.com/raruidol/ArgumentRelationMining
Model: https://huggingface.co/raruidol/ArgumentRelation

Argument Mining (III): Limitations

- Annotations are done for specific computational argumentation tasks. (e.g., argument segmentation, relation identification, argument summarisation, etc.)

- Limited size of existing corpora due to the elevated complexity of annotating new data.

- Unavailable annotations of complete argumentative discourses and debates.

- Non-professional and inexpert sources and domains (w.r.t. purely argumentative purposes):
  - Political debates: fallacious reasoning, voters' persuasion, etc.
  - Online forums: anonymous identity, lack of filtering and rules, etc.

→ A complete automatic analysis of argumentative discourse is not feasible.

The *VivesDebate* Corpus (I)

VivesDebate: A New Annotated Corpus of Argumentation in a Debate Tournament[2].

- https://zenodo.org/record/6531487
- Debate Tournament: Xarxa Vives d'universitats.
- 29 debates; 139,756 words.
- Argumentative and structural annotations.
  - → 7,810 Argumentative discourse Units (ADUs)
  - → 12,653 Inferences (RA)
  - → 1,558 Conflicts (CA)
  - → 747 Rephrases (MA)
- Debates evaluated by impartial experts:
  - → Score: Thesis solidity; Argumentation Quality; Adaptability
- Complete annotation of debates.
- Languages: CAT (original), ES, EN

The *VivesDebate* Corpus (II)

Introduction (INTRO) 8 min	Proponent team: Presentation of its main lines of argument	4 min
	Opponent team: Presentation of its main lines of argument	4 min
Argumentation (ARG1) 12 min	Proponent team: Development of its main arguments	6 min
	Opponent team: Development of its main arguments and rebuttal	6 min
Argumentation (ARG2) 12 min	Proponent team: Reinforcement of its main arguments and rebuttal	6 min
	Opponent team: Reinforcement of its main arguments and rebuttal	6 min
Conclusion (CONC) 8 min	Opponent team: Conclusion of the debate	4 min
	Proponent team: Conclusion of the debate	4 min

The *VivesDebate* Corpus (III)

Table 2: Structure of the *VivesDebate* annotations.

ID	Phase	Arg. Number	Stance	ADU_CAT, ADU_ES, ADU_EN	Related ID	Relation Type
1	INTRO		FAVOUR	*quan mireu aquí què veieu* (*cuando miráis aquí qué veis*) (*when you look here what do you see*)		
2	INTRO		FAVOUR	*cinquanta euros* (*cincuenta euros*) (*fifty euros*)	1	RA
...	...	...	...	...	...	...
45	ARG1	1	FAVOUR	*i si som guanyadors podem comprar tot allò que desitgem* (*y si somos ganadores podemos comprar lo que deseemos*) (*and if we are winners we can buy whatever we want*)	44	RA
46	ARG1	1	FAVOUR	*és el model que està imperant en la gestació subrogada* (*es el modelo que impera en la gestación subrogada*) (*is the prevailing surrogacy model*)	43;44;45	RA
...	...	...	...	...	...	...
144	ARG1	3	AGAINST	*per supossat que no* (*por supuesto que no*) (*of course not*)	143	CA
145	ARG1	3	AGAINST	*no és que vullguem que hi haja més xiquets* (*no es que queramos que haya más niños*) (*not that we want there to be more children*)	140	RA
146	ARG1	3	AGAINST	*sinó tot el contrari* (*sino todo lo contrario*) (*quite the contrary*)	145	MA
...	...	...	...	...	...	...

The *VivesDebate* Corpus (IV)

Table 3: Structure of the *VivesDebate* evaluation file.

Debate	Stance	Score	Thesis Solidity	Argument Quality	Adaptability
Debate1	Favour	3.32	3.25	3.37	3.33
Debate1	Against	3.29	3.33	3.18	3.38
Debate2	Favour	3.41	3.5	3.33	3.42
Debate2	Against	3.43	3.17	3.58	3.58
. . .	. . .	. . .	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .	. . .
Debate28	Favour	2.75	3.25	2.75	2.17
Debate28	Against	2.36	2.77	2.25	2.00

The *VivesDebate* Corpus (V)

Table 4: Comparison of computational argumentation corpora. (∗) Automatically translated languages.

Identifier	Domain	Tasks	Language	Size	Annotation Ratio
Persuasive Essays	Online Forum	AM	EN	34,917 (W)	388 *w/d*
Microtexts	Controlled Experiment	AM	EN+DE	576 (S)	5 *s/d*
DMC	Academic+Online+Professional	AM+AA	EN	18,628 (W)	144 *w/d*
CDCP	Online Forum	AM+AE	EN	4,931 (S)	6.7 *s/d*
GPR-KB-55	Competitive Debate	AG	EN	12,431 (S)	41 *w/d*
US2016	Political+Online	AM+AA	EN	97,999 (W)	189 *w/d*
US2016TV	Political	AM+AA	EN	58,900 (W)	492 *w/d*
US2016Reddit	Online Forum	AM+AA	EN	39,099 (W)	137 *w/d*
DebateSum	Competitive Debate	AS	EN	101M (W)	520 *w/d*
ReCAP	Political	AM+AA	DE+*EN*(∗)	16,700 (W)	150 *w/d*
VivesDebate	Competitive Debate	**AM+AA+AE+AG/AS**	CAT+*ES*(∗)+*EN*(∗)	139,756 (W)	**4819 *w/d***

Computational Argumentation Tasks: Argument Mining (AM), Argument analysis (AA), Argument Evaluation (AE), Argument Generation (AG), Argument Summarising (AS)
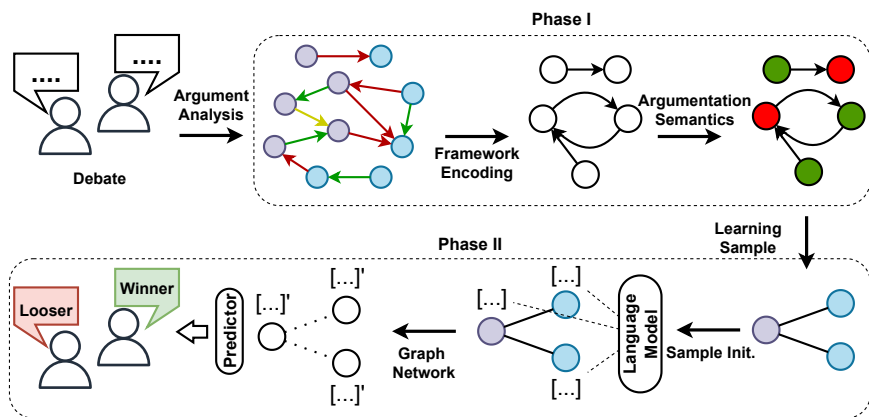
Available in Zenodo: https://zenodo.org/record/6531487

The *VivesDebate* Corpus: Advantages (VII)

- One of the largest publicly available resources annotated with relevant argumentative propositions and relations
- Quality: professional debate (well-supported arguments, no demagogy), objectively evaluated by an impartial and expert jury
- Versatility: useful for a wide range of argumentative tasks such as argumentative language modelling, the automatic identification of ADUs (i.e., Argument Mining), the elaboration and analysis of complex argument graphs (i.e., Argument Analysis), the automatic evaluation of arguments and argumentative reasoning (i.e., Argument Evaluation), and the automatic generation of argument summaries (i.e., Argument Generation/Argument Summarising)
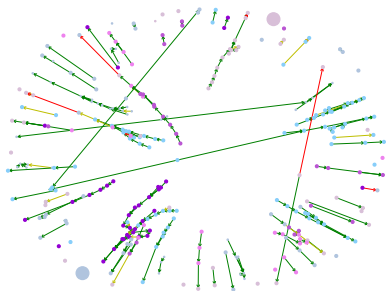- Multilingual: Catalan (original), Spanish, English

Automatic Debate Evaluation with Argumentation Semantics and Natural Language Argument Graph Networks[3].
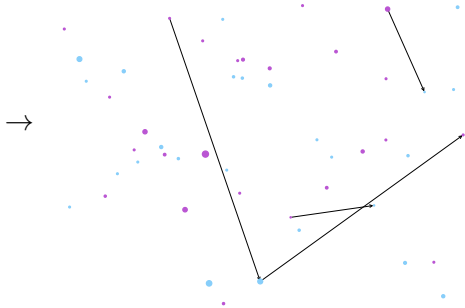
[3] Ramon Ruiz-Dolz, Stella Heras, and Ana García-Fornes. "Automatic Debate Evaluation with Argumentation Semantics and Natural Language Argument Graph Networks". In: *arXiv preprint arXiv:2203.14647* (2022).

Automatic Evaluation of Argumentative Debates (II)

Argumentative analysis graph:

Argumentation Framework:

$\rightarrow$

Definition (Abstract Argumentation Framework)

An Abstract Argumentation Framework (AAF) is a tuple $AAF = <A, R>$ where: A is a finite set of arguments, and R is the attack relation on A such as $A \times A \rightarrow R$.

Automatic Evaluation of Argumentative Debates (III)

Argumentation Semantics:
$\rightarrow$ Logical rules to determine argument acceptability.

Definition (Conflict-free)

Let $AF = <A, R>$ be an argumentation framework and $Args \subseteq A$. The set of arguments $Args$ is conflict-free *iff* $\neg\exists\alpha_i, \alpha_j \in Args : (\alpha_i, \alpha_j) \in R$.
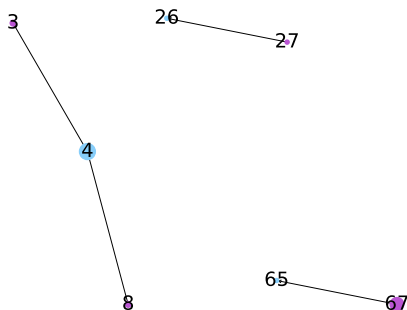
Definition (Admissible)

Let $AF = <A, R>$ be an argumentation framework and $Args \subseteq A$. The set of arguments $Args$ is admissible *iff* $Args$ is conflict-free, and $\forall\alpha_i \in Args, \neg\exists\alpha_k \in A : (\alpha_k, \alpha_i) \in R$ and $(\alpha_i, \alpha_k) \notin R$, or $\neg\exists\alpha_j \in Args : (\alpha_j, \alpha_k) \in R$

$\rightarrow$ **Naïve** (conflict-free) vs. **Preferred** (admissible) *maximal* semantics.

Argumentation Semantics:

$\mathcal{E}_1$: [3, 8, 26, 65] $\mathcal{E}_3$: [3, 8, 27, 65]

$\mathcal{E}_2$: [3, 8, 26, 67] $\mathcal{E}_4$: [3, 8, 27, 67]

+ All non conflicting arguments.

Automatic Evaluation of Argumentative Debates (V)

| Experiment | Train | | Eval. Metrics | |
Model	D	S	Acc.	Macro-F1
Naïve-GN	23	369	**0.72**	**0.48**
Preferred-GN	23	26	0.40	0.40
Naïve-ATB	-	-	0.16	0.14
Preferred-ATB	-	-	0.20	0.16
LMB	23	23	0.60	0.37
RB	-	-	0.48	0.33

Table 5: Accuracy and Macro-F1 results of the automatic debate evaluation task. D and S indicate the number of debates and learning samples respectively used in the Train data partition in our experiments.

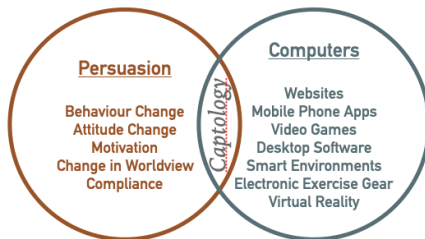Code: https://github.com/raruidol/ArgumentEvaluation

Automatic Evaluation of Argumentative Debates (VI)

- Naïve-GN Graph Network
- Preferred-GN Graph Network
- Four baselines:
  - Random Baseline (RB): assigns randomly a class to each extension
  - Naïve Argumentation Theory Baseline (Naïve-ATB): classifies each Naïve extension by counting the majority number of acceptable arguments belonging to each stance
  - Preferred Argumentation Theory Baseline (Preferred-ATB): classifies each Preferred extension by counting the majority number of acceptable arguments belonging to each stance
  - Language Modelling Baseline (LMB): implemented ignoring the Phase I and training the GN directly over the whole argumentative analysis graphs.

Conflict-free semantics produce a higher number of acceptable extensions from each AF. This helped to improve the learning power of the argument evaluation task in a similar way to that achieved by data augmentation.

Outline

" Captology: Computers as Persuasive Technologies

Persuasion

Behaviour Change
Attitude Change
Motivation
Change in Worldview
Compliance

Computers

Websites
Mobile Phone Apps
Video Games
Desktop Software
Smart Environments
Electronic Exercise Gear
Virtual Reality

Captology

-B.J. Fogg, Stanford Persuasive Tech Lab

- Persuasion is one of the most common forms of (computational) argumentation dialogues, but ...
- Argumentation is not central to the current manifestations of persuasion technologies.
- The arguments for good behaviour seem either to be assumed before the user accesses the persuasion technology (e.g. when using diaries, or receiving email reminders), or arguments are provided implicitly in the persuasion technology (e.g. through provision of information, or through game playing). [A. Hunter. https://www.computationalpersuasion.com/]

Towards an Argumentation System for Assisting Users with Privacy Management in Online Social Networks[4].

Definition (Argumentation Framework for Online Social Networks)

We define an argumentation framework for online social networks as a tuple $AFOSN = <A, R, P, \tau_p>$ where: A is a set of n arguments $[\alpha_1, \ldots, \alpha_n]$; R is the attack relation on A such as $A \times A \to R$; P is the list of e profiles involved in an argumentation process $[p_1, \ldots, p_e]$; and τ_p is a function $A \times P \to [0, \ldots, 1]$ that determines the score of an argument α for a given profile p.

---

[4]Ramon Ruiz-Dolz et al. "Towards an Argumentation System for Assisting Users with Privacy Management in Online Social Networks.". In: *CMNA@ PERSUASIVE*. 2019, pp. 17–28.
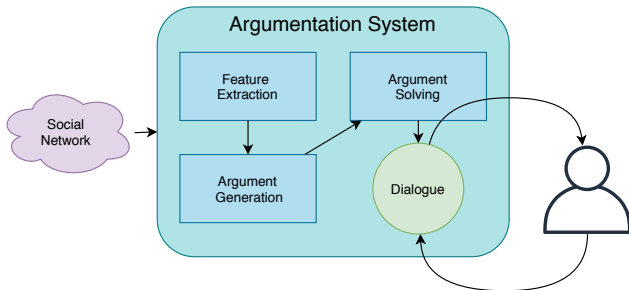
Argumentation in Online Social Networks (II)

Figure 1: `https://raruidol.github.io/videos/demo.mov`

"On the Prevention of Privacy Threats: How Can We Persuade Our Social Network Users?" In: arXiv preprint arXiv:2104.10004 (2021).
→ Users with specific profiles (personality traits and OSN usage trends) were more effectively persuaded with specific types of arguments (content, privacy, risk, trust).

A Qualitative Analysis of the Persuasive Properties of Argumentation Schemes[5].

- Not much attention on the domain-agnostic relation between persuasive strategies and the argumentative patterns of human reasoning.
- Improve our understanding of the persuasive properties of argumentative reasoning.
- Provide a domain-independent approach for argument-based persuasion.
- Support the evolution of new computational argumentation systems by improving the interaction with human users.

---

[5]Ramon Ruiz-Dolz et al. "A Qualitative Analysis of the Persuasive Properties of Argumentation Schemes.". In: *UMAP2022*. 2022, In press.

Argumentation Schemes and Cialdini's Principles of Persuasion(II)

$\rightarrow$ Study the relationship between Cialdini's principles of persuasion and argumentation schemes (stereotyped patterns of human reasoning).

- **Cialdini's Principles of Persuasion:**
  - *Reciprocity*
  - *Authority*
  - *Commitment*
  - *Liking*
  - *Social Proof*
  - *Scarcity*

- **Argumentation Schemes:**
  - *Argument from Popular Opinion* (AFPO)
  - *Argument from Popular Practice* (AFPP)
  - *Argument from Position to Know* (AFPK)
  - *Argument from Expert Opinion* (AFEO)
  - *Argument from Commitment* (AFCM)
  - *Argumentation from Values* (AFVL)
  - *Argument from Practical Reasoning* (AFPR)
  - *Argument from Waste* (AFWS)
  - *Argument from Sunk Costs* (AFSC)
  - *Argument from Threat* (AFTH)
  - *Argument from Cause to Effect* (AFCE)
  - *Argument from Rules* (AFRL)

Argumentation Schemes and Cialdini's Principles of Persuasion(III)

$\rightarrow$ Experiment in the platform *Prolific* with 117 participants.

4 Topics of debate

- Should COVID-19 Coronavirus vaccination be mandatory?
- Should you take care of your physical appearance to achieve personal and professional success?
- Should euthanasia be legalised?
- Should you do intermittent fasting to lose weight?

2 Stances

- In Favour / Against

96 Arguments

Argumentation Schemes and Cialdini's Principles of Persuasion(IV)

Instance of an argumentation scheme

Most people think that through mass vaccination it would be possible to reduce ...

Possible options

- *Reciprocity: "People feel obliged to return to others ...*
- *Authority: "People accept the opinions of knowledgeable experts ...*
- *Consistency: "People like to be consistent with the things they ...*
- *Sympathy: "People prefer to accept the opinions of someone ...*
- *Consensus: "Especially when we are not sure, people look to ...*
- *Scarcity: "People want more of what they can have less of ...*
- *Other*

A Qualitative Analysis of the Persuasive Properties of Argumentation Schemes[6].

Table 6: Argumentation schemes' principles of persuasion.

Argumentation Scheme	Cialdini's Principle of Persuasion
AFPK, AFEO	Authority
AFCM, AFSC	Commitment
AFPO, AFPP	Social Proof
AFVL	*Commitment**
AFTH	*Reciprocity**
AFRL	*Authority-Commitment**
AFPR, AFWS, AFCE	None

[6]Ruiz-Dolz et al., "A Qualitative Analysis of the Persuasive Properties of Argumentation Schemes."

Outline

Ongoing Research

Definition and Evaluation of a Generic Argument-based Persuasive Framework
$\rightarrow$ Considering all the previous findings, how can we define a user-tailored argumentation system and improve its persuasiveness?

Emotions and Argumentation
$\rightarrow$ How can emotions influence the human perception of reasoning patterns (i.e., argumentation schemes) and its persuasiveness? $\rightarrow$ Study the relationship between reasoning patterns and emotions. $\rightarrow$ Study the relationship between persuasion principles and emotions.

$\rightarrow$ New experiment: present arguments (argumentation schemes) and statements formulated with Cialdini's Principles of persuasion to subjects, while monitoring their emotions (with hardware)