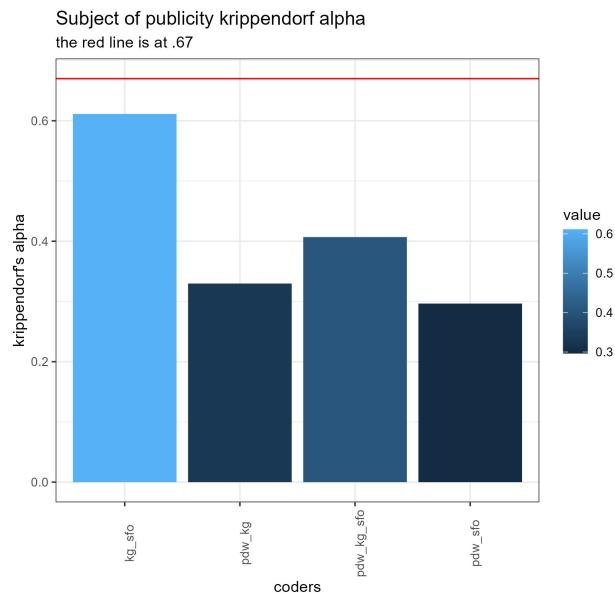


Qualitative analysis of intercoder reliability test

Sina Furkan Özdemir

2/23/2022

subject of publicity (SoP) coding differences:



Sina and Kristine:

There are total of 24 dissonant coding between me and Kristine out of 100 tweets. Upon closer examination there are three core coding differences in my and Kristine's coding about subject of publicity. These are *tacit agency in the message, retweets that include plural personal pronouns (i.e we, our)*, and *the interpretation of agency of mentioned actors*.

The first systematic difference is the threshold I and Kristine uses to attribute agency to mentioned actors in a tweet. This case is illustrated by figure 1 and 2. Kristine attributed agency to @EU_Env and @RanInitiative in these tweets while coding subject, whereas Sina did not. In some other cases, the employed threshold of agency is reversed among coders where Sina attributed agency but Kristine did not. Such differences amount up to 37.5% of the dissonance among the coders.

The second one is assuming agency in the tweets. As the figure 3 below illustrates, there are tweets without explicit mention of an agent behind the output. In most cases, I tend to code these tweets with subject of publicity: none while Kristine seems more prone to attributing agency to the tweeting account. This accounts for 29.1666667% of the difference in the coding. We decided that we will allow some level of interpretation of agency even if it is not explicitly mentioned in the tweet in our previous meetings. However, this decision was made in the light of tweets such as the figure 4, where the agency is almost explicit (fully explicitly



Hauts-de-France EU

@hautsdefranceEU



[#EUDIALOGUES]

Des questions sur le [#PacteVert](#), le plan d'action en faveur de l'économie circulaire ou la stratégie en faveur de la [#biodiversité](#) ? Le Commissaire européen environnement, océans et pêche répond à vos questions le 24 avril ! Envoyez vos questions à [@EU_env](#) !
[twitter.com/EU_ENV/status/...](https://twitter.com/EU_ENV/status/)

EU Environment [@EU_ENV](#)

The #EUGreenDeal, #CircularEconomy & #EUBiodiversity strategies are key for the green recovery for Europe 🌱

Send us your questions to Virginijus Sinkevičius [@VSinkevicius](#), Commissioner for Environment, Oceans & Fisheries
=> #EUDIALOGUES LIVE chat on Friday 24 April, 2pm CET



♡ 4 9:04 AM - Apr 21, 2020



See Hauts-de-France EU's other Tweets





New Zealand Mission to the EU ✅



@NZinEU

This [@EUHomeAffairs' @RanEurope](#) initiative in preventing #violentextremism is exactly what our collective #ChristchurchCall is all about. Digital youth workers & communications experts interested in the meeting taking place this March can apply by 31 Jan: bit.ly/2Rkzu6A twitter.com/RANEurope/stat...

Radicalisation Awareness Network – RAN ✅ @RANEurope

📢 Call for participants

Are you a (digital) youth worker or a communications expert interested in preventing #violentextremism online? Join the RAN YF&C – C&N meeting in Brussels and deepen your knowledge. Apply by ⌚ 31 Jan. 👉
bit.ly/2Rkzu6A



Heart 5 11:50 AM - Jan 24, 2020



See New Zealand Mission to the EU's other Tweets



Figure 2: 2

when the multi-media is taken into account). However, the decision is not formalized in the codebook with a boundary on the interpretation.

The last difference is our approach to retweeted tweets with personal pronoun. There are several tweets where cascading mention of actors create a very complex agency reporting structure as illustrated in the figure 5. In most cases, these manifest themselves as an account retweeting another account's tweet which includes a personal pronoun. In some cases, agents in the retweeted tweet includes the retweeter but not always. Kristine and I approach interprets the agency structure in these tweets differently. This difference accounts for 8.333333% of the dissonance in the subject of publicity coding. Overall, I tends to include the retweeting account in the set of agents while coding the retweet, thus coding them as **compound** while Kristine tends to exclude the retweeter from the set thus coding the subject of publicity as **other actors**.

Besides these systematic differences, most of the dissonance in the subject of publicity between coders are due to human error. For example, I equated @EU_GNSS to @EU4Space in the case of figure 6 as if they are the same accounts while obviously they are not the same agency.

Possible solution It is hard to identify a systematic difference in the results however. Only in dissonance 1 there seems to be a coding pattern as illustrated by the table-1 below:

tweet_id	sfo	kg
x1214489140466716673	5	2
x1235245669989986304	5	3
x1250722806591086593	5	2
x1222441176160260097	5	2
x1288373984610967554	5	2
x1230168541653020673	5	1
x1225346800904740864	5	1

It is possible to sort type 1 dissonance if I revise the tweets I coded as subject of publicity: “other” if Kristine and I can agree on the threshold of agency.

Sina and Pieter

There are total of 45 differences between my and Pieter's coding out of 100. In similar vein, the difference between Pieter's and my coding are due to three key dissonance. These are namely *threshold to attribute agency*, *retweets that include plural personal pronouns (i.e we, our)* and *application of coding rules*. Very much like the coding difference between my coding and Kristine's, it seems like I have a higher threshold to attribute agency to mentioned accounts. This dissonance accounts for 20% of the coding difference. Figure 7 illustrates this point.

In this example, I coded attributed agency to only @EUHomeAffairs and coded SoP as self because the tweet says “@EUHomeAffair's @RanEurope initiative”. Pieter, on the other hand, followed a more stringent approach and attributed agency to all mentioned accounts in the tweet, thus coded it as compound.

The dissonance 2 regarding how to approach retweets with personal pronouns seems to be another regular difference between my coding and Pieter. As illustrated by figure 8, I seem to perceive these messages as retweeting account talking about other actors practices and code it as such but Pieter seems to apply the exception rule for “compound” category.

Last type of dissonance is the inconsistent application of coding rules which amounts up to $r(5/45)*100\%$ of the coding differences. This broad category of dissonance includes two key differences. The First and foremost is about how we treat the retweets. It seems Pieter miss the retweets and code the tweets as if it is primary tweet such as the following figure.



European Commission



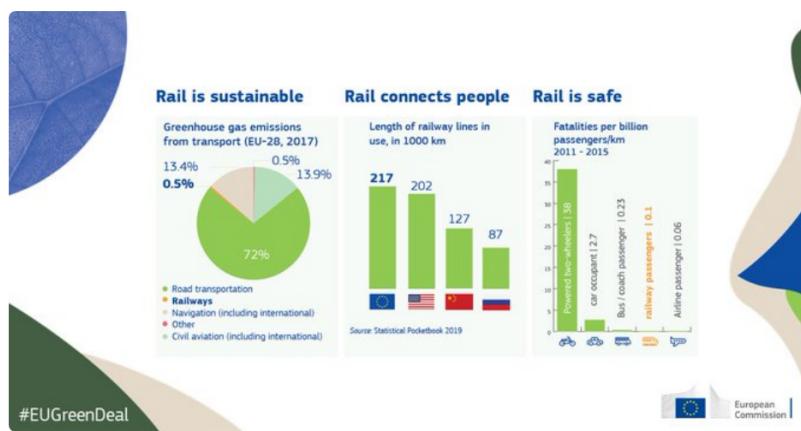
@EU_Commission · Mar 4, 2020



Replies to @EU_Commission

Trains genuinely connect Europeans. They're one of our greenest and safest modes of transport.

Rail is the only mode of transport to have almost continuously reduced its CO₂ emissions since 1990, at the same time as it increased transport volumes. #EUYearofRail #EUGreenDeal



#EUGreenDeal



European Commission



European Commission



@EU_Commission

The #EUYearofRail will highlight benefits for people, economy and climate.

The focus will be on the remaining challenges to create a true Single European Rail Area without borders.

More here → europa.eu/ljd36Vq #EUGreenDeal



#EUGreenDeal

WHAT IS IT ABOUT?

- Reaching the European Green Deal objectives
- Promoting attractive and sustainable way to connect people and businesses
- Complementing the future Strategy for Sustainable and Smart Mobility
- Highlighting other key initiatives
 - Shift2Rail
 - Revision of the TEN-T Regulation
 - Fourth Railway Package



50 5:48 PM - Mar 4, 2020



19 people are talking about this





Kadri Simson



@KadriSimson

Excellent meeting with [@aziz_rabbah](#), Moroccan Minister of Energy. We agreed to strengthen our cooperation in the perspective of EU-Africa relations. Great to hear about Morocco's activities and initiatives on electricity, renewables, energy efficiency and hydrogen.



20 5:59 PM - Dec 6, 2019



[See Kadri Simson's other Tweets](#)





EU Climate Action ✅

@EUClimateAction



#ClimateAction starts with each & every one of us!

That's why we're excited to start the #EUClimatePact together with all of you!

We want to bring everyone together, join forces & build action for a greener future 🌱

Stay tuned for the announcement on 4/3! europa.eu/lvf39uu



187 2:40 PM - Feb 17, 2020



124 people are talking about this



Figure 5: 5



Eurisy
@Eurisy1



New #Horizon #EGNSS market uptake 2019-2020 Call. Dealing with the development of new innovative #applications including a new topic tailored to #publicauthorities.

! Deadline: 5 March ! bit.ly/2q3RE1i @EU_GNSS



European
Global Navigation
Satellite Systems
Agency



3 8:50 AM - Feb 20, 2020



See Eurisy's other Tweets



Figure 6: 6



New Zealand Mission to the EU ✅

@NZinEU



This [@EUHomeAffairs' @RanEurope](#) initiative in preventing #violentextremism is exactly what our collective #ChristchurchCall is all about. Digital youth workers & communications experts interested in the meeting taking place this March can apply by 31 Jan: bit.ly/2Rkzu6A twitter.com/RANEurope/stat...

Radicalisation Awareness Network – RAN ✅ @RANEurope

📢 Call for participants

Are you a (digital) youth worker or a communications expert interested in preventing #violentextremism online? Join the RAN YF&C – C&N meeting in Brussels and deepen your knowledge. Apply by ⌚ 31 Jan. 👉
bit.ly/2Rkzu6A



♡ 5 11:50 AM - Jan 24, 2020



👤 See New Zealand Mission to the EU's other Tweets



Figure 7: 7

A global threat needs a global response. Today I urged #G7 Health Ministers to work together on #COVID19 vaccines development and join our efforts for a 🌎 pledge on 4 May to fund the #coronavirus challenge.

62 2:10 PM - Apr 16, 2020

28 people are talking about this

Figure 8: 8

Second difference is Pieter's interpretation of the use of plural personal pronouns in tweets by institutional accounts. He seems to interpret agency as "self" whereas I tend to apply the exception rule more often as illustrated by the figure below.

Beyond these three types of dissonance, the rest of the differences seem idiosyncratic human error.

Possible solutions The following cross tabulation illustrates the coding differences. The most common coding pattern where it is sfo:"other actors" - pdw:"compound" is due to the fact that I missed the mention of certain actors in the tweet (i.e human error). The second most common coding pattern where sfo:"none" - pdw:"other actors" is partially due to dissonance type 1 and partially due to human error.

'summarise()' has grouped output by 'sfo'. You can override using the '.groups' argument.

sfo	pdw	n	perc_shar
1	2	2	4.4
1	3	3	6.7
1	5	2	4.4
2	1	4	8.9
2	3	10	22.2
2	5	5	11.1
3	1	5	11.1
3	2	4	8.9
3	5	1	2.2
5	1	1	2.2
5	2	8	17.8

At this point, there are a few possible solutions to remedy these dissonances. For example, I could introduce an attention check question to the questionnaire that forces the coder to pay attention to whether it is a retweet or not. Similarly, a set of clarification rules about how to deal with SoP in retweets in the codebook may be useful. However, considering the fact that Pieter is almost done with his sample, these would not help with the quality of the extant dataset. So, I concur with Pieter's suggestion to remove his coding from

INHOPE  **INHOPE** 
@INHOPE_PR



Thank you to all the attendees of the INHOPE and @EU_Commission Focus Group on the use of #AI to combat #CSAM online. Together we can build a roadmap for future use of AI! #cybersecurity
#ArtificialIntelligence #SID2020 @safeinternetday #SaferInternetDay



♡ 25 6:02 PM - Feb 12, 2020



See INHOPE's other Tweets



Figure 9: 9



Wetlands are disappearing three times faster than forests.

Learn why we need to save these vital ecosystems.

ow.ly/AYFm50xCTyP @IUCN_Water #Wetlands



349 10:30 AM - Jul 19, 2020



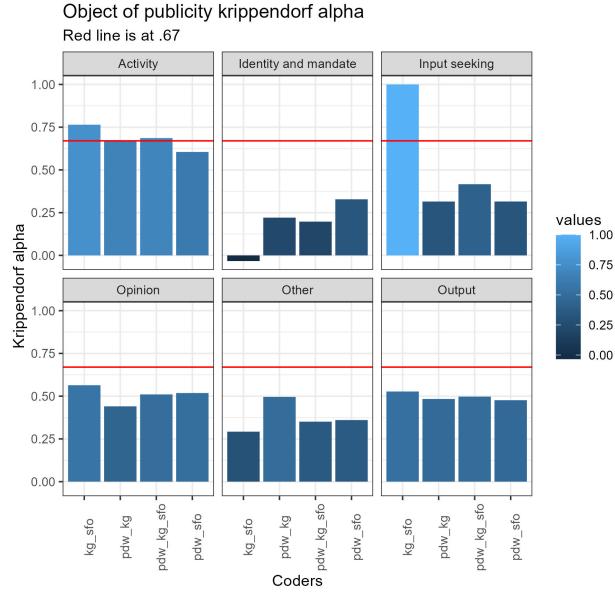
236 people are talking about this



Figure 10: 10

the dataset and recode them myself. Therefore, I will only focus on the coding differences between Kristine and I.

Object of Publicity differences



Identity and mandate

Upon closer examination, there is only 8 out of 100 tweets differently coded between my and Kristine's coding. All of them seem to be due to different interpretation of the category. Figure below illustrates these differences.

Here, I interpreted ”*@EUEic will make a huge impact by emulating positive features of traditional VCs & investing in risky research needed to turn scientific breakthroughs into viable products.*” but Kristine did not. I think these differences are mainly due to the fact that the definition for the category is changed recently and still quite vague. A revision of the definition and coding as well as a repeated test of intercoder reliability of this category alone should resolve the problem. A more pragmatic approach would be to either drop this variable as suggested by Pieter and Kristine in the last meeting. Since identity and mandate is only one indicator of broader theoretical concept of one-way communication strategy, we would still have enough to proceed with the analysis.

Output

There are total of 21 dissonant codings of output category between my and Kristine's coding. There seems to be several key patterns causing the difference. The first is the tweets that are in and of themselves are output such as public service announcements as illustrated by the figure below. Kristine coded these as output while I mostly coded this as “other” in the object of publicity categories.

The second source of dissonance is potential borderline cases between output, identity, opinion and activity. Following examples through figure 13 to 15 illustrate this point. This probably due to the fact that a) these cases are borderline and b) the definition of output category is not clear enough. Since I still have quite a bit of tweets to code, agreeing on a practice should improve the situation for the rest of the dataset.

The last possible source of dissonance is the lower threshold to assume a tweet to be an output message. As illustrated by the figure 16 below, I didn't interpret the tweet as an output message but Kristine did. Similar



Jean-Eric Paquet

@JEPaquetEU



NEWS: [@EU_Commission](#) is planning a €3.5bn fund for early stage tech! [@EUEic](#) will make a huge impact by emulating positive features of traditional VCs & investing in risky research needed to turn scientific breakthroughs into viable products. More here bloomberg.com/news/articles/...



EU Plans \$3.9 Billion Fund for Startups in 'Valley of Death'

The European Union is planning a 3.5 billion-euro (\$3.9 billion) fund that will invest in early stage technology in an effort to increase the pipeline of bloomberg.com

158 12:39 PM - Nov 25, 2019



98 people are talking about this



Figure 11: 11



EESC Agriculture RuralDev Environment

@EESC_NAT



📢 Do you know about the 2020 strategy on #EUBiodiversity?
Today's conference looks at how to halt the loss of #biodiversity
by delivering on new standards in

📈 trade

🏗️ industry

🌿 agriculture

เศรษconomy

Watch live 🎙 twitter.com/i/broadcasts/1...



Watch live our conference on #EUBiodiversity!

EESC Agriculture RuralDev Environment @EESC_NAT

16 9:34 AM - Jan 29, 2020



15 people are talking about this



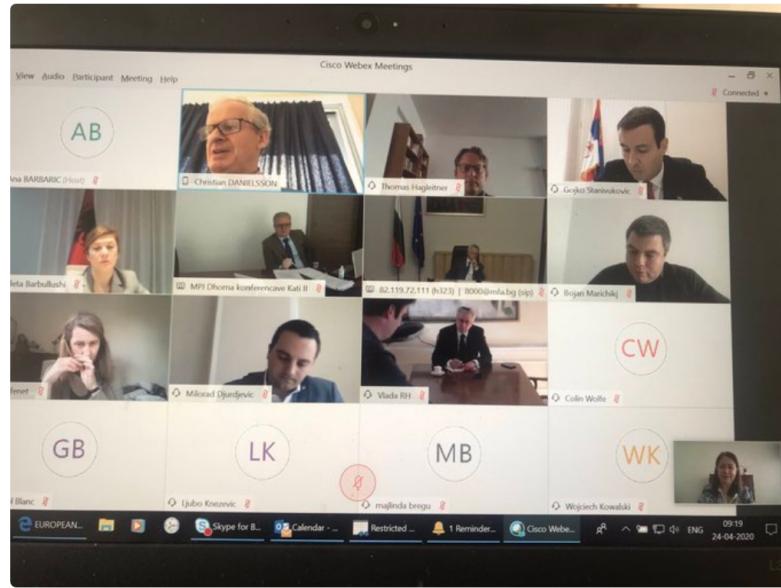
Figure 12: 12



Genoveva Ruiz Calavera
@CalaveraRuiz

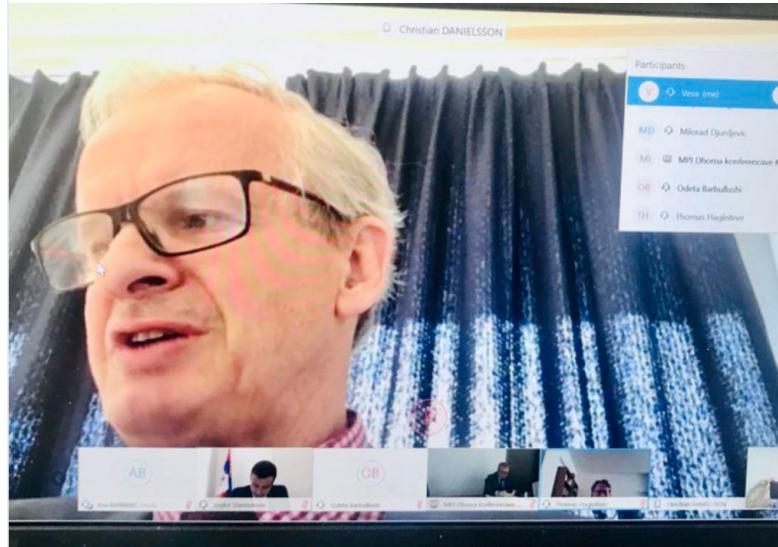


Thanks @DanielssonEU @eu_near for hosting the VTC today. Sherpas of Prime Ministers of the Western Balkans, Regional organisations: @rccint, @CEFTA_, @EUWB6_TCT, 🇪🇺 officials and @EU2020HR working together for the benefit of all 🇧🇦 🇩🇪 🇮🇹 🇷🇸 🇳ශ 🇮ත citizens! #StrongerTogether [twitter.com/DanielssonEU/s...](https://twitter.com/DanielssonEU/status/1253441071000000000)



Christian Danielsson @DanielssonEU

Excellent exchange💻📞 with Sherpas of Prime Ministers of the #WesternBalkans in advance of #EUCO🇪🇺 meeting with region on 06 May. #StrongerTogether @eu_near



33 3:31 PM - Apr 24, 2020



16 people are talking about this >

Figure 13: 13
16



marco marinucci
@mmarinucci



the EU green-deal presented [@UCBerkeley](#) by [@JEPaquetEU](#) - could that be a GameChanger for EU? [#climatechange](#) [#SEC2SV](#) cc [@isidrolaso](#) [@MindTheBridge](#)



12 9:45 PM - Feb 10, 2020



[See marco marinucci's other Tweets](#)



Figure 14: 14



Valdis Dombrovskis ✅ @VDombovskis · Apr 22, 2020



The 🇪🇺 EU wants to support its partner countries in dealing with #coronavirus economic impact - @EU_Commission plans special use of macro-financial assistance loans to help them address immediate financing needs, like healthcare 👏 bit.ly/2VtrSjX



Valdis Dombrovskis ✅

@VDombovskis

As part of the 🇪🇺 EU's global response to the coronavirus pandemic, we need to help our neighbouring countries to cushion the worst of its economic impact and keep the entire region stable. #coronavirus

Supporting our neighbours is essential during this time of crisis to keep the entire region stable. As part of the EU's global response to the coronavirus pandemic, we need to help our neighbouring countries to cushion the worst of its economic impact. These 'crisis MFA programmes' will assist 10 countries in ensuring macro-economic stability and protecting their people and companies during the crisis.

Valdis Dombrovskis
Executive Vice-President for
An Economy that Works for People

European Commission |

13 11:19 AM - Apr 22, 2020 · European commission



See Valdis Dombrovskis's other Tweets



to the last source, making a decision on how to proceed with such tweets should be sufficient to improve the data quality for the rest of the dataset.



EP Research Service ✅

@EP_ThinkTank



#WorldCancerDay - what's the EU plan to beat cancer?
europarl.europa.eu/thinktank/en/d... @EP_Environment @MAC_MEPs
@VTrillet_Lenoir @loucas_fourlas @Miriamdalli @CancerLeagues
@SKyriakidesEU #EUCancerplan



12 1:54 PM - Feb 4, 2020



See EP Research Service's other Tweets



Figure 16: 16