

Electronic Nose for Odor Classification of Grains

T. BÖRJESSON,^{1,4} T. EKLÖV,² A. JONSSON,¹ H. SUNDGREN,² and J. SCHNÜRER³

ABSTRACT

Cereal Chem. 73(4):457-461

An electronic nose was used to classify grain samples based on their smell and to predict the degree of moldy/musty odor. A total of 235 samples of wheat, barley and oats, which had been odor classified by at least two grain inspectors, were used. Headspace samples from heated grain were pumped through chambers containing metal oxide semiconductor field effect transistor (MOSFET) sensors, SnO₂ semiconductors and an infrared detector monitoring CO₂. The sensor signals were evaluated with a pattern-recognition software program based on artificial neural

networks. The samples were divided into either the four classes moldy/musty, acid/sour, burnt, or normal or the two classes good and bad according to the inspectors descriptions. They were also assigned a score describing their intensity of moldy/musty odor. The electronic nose correctly classified ≈75% of the samples when using the four-class system and ≈90% when using the two-class system. These values exceeded the corresponding percentages of agreement between two grain inspectors classifying the grain.

In Sweden, as well as in many other countries, grains are checked for off-odors upon delivery at granaries. Off-odors make grains and grain products less palatable and are often indicative of past or ongoing microbial deterioration. Thus off-odor characterization offers a potential way for quickly and cheaply assessing batches of grain to determine whether they should be accepted for human or animal consumption, used for other purposes, or rejected.

Odors are described as either normal, musty, moldy, acid, sour, burnt, or foreign, and the intensities of off-odors are given as weak, pronounced, or strong (Statute Book 1991). Because of the cool climate in Sweden, insect infestation is unusual, and thus, insect odor is not present among the off-odors that are checked.

However, the procedure suffers from at least two drawbacks. The first drawback is lack of objectivity. Inevitably, there are disagreements between human individuals in terms of how they perceive types and intensities of odors. For example, Stetter (1992) studied the classification of 12 samples of wheat into the five odor categories, normal, insect, musty, foreign, or sour, by four inspectors. Unanimous agreement was obtained for only four of the samples. However, when all off-odors (insect, musty, and foreign) were lumped together into one category, unanimous agreement as to whether the samples were normal or off-odorless was obtained for eight samples. The second drawback is the health aspect. Inhalation of mold spores from damaged grain can induce allergic reactions (Rylander 1986), and exposure to fungal volatile metabolites can cause various disease symptoms (Samson 1985).

Thus, it would be advantageous to develop an instrumental replacement for the inspector. Compounds that cause off-odors in grains can be measured using gas chromatography and mass spectrometry (Börjesson et al 1993). However, the high costs and limited time available at granaries make this approach unsuitable. By contrast, using an array of nonspecific sensors coupled to a pattern-recognition routine, should make it possible to screen grain quickly and cheaply. Furthermore, this procedure mimics the way odors are perceived in humans and other animals.

Artificial neural networks (ANN) show good promise as data-processing tools (Stetter 1992). During the learning phase of this approach, a training set of response patterns is first presented to the ANN along with respective class affiliations. Performance is then measured as the percentage of odors classified correctly when presenting a test set of new patterns to the ANN. Stetter et al (1993) recently used an ANN to correlate odor descriptions in wheat to output from an array of electrochemical sensors. They found that the ANN correctly classified 83% of the wheat samples when using three odor categories. However, they did not use true unknown samples; performance was tested with samples included in the training set. Furthermore, the technique used, which includes the trapping and purging of volatile compounds, would probably be too time-consuming for grain-reception applications.

Sensors for an electronic nose should be sensitive to aroma compounds, which are mainly hydrophobic compounds with molecular mass of 18–300 Da. Furthermore, the various types of sensors used should differ in terms of the types of compounds to which they respond (Bartlett and Gardner 1992).

SnO₂ semiconductor sensors are sensitive to combustible gases, of which many are odorless. Such sensors come in many types, but the responses to mixtures of aroma compounds from foods are highly correlated (Aishima 1991). In view of the great similarity in the compounds they responded to, they should be combined with other types of sensors. One suitable type is the metal oxide semiconductor field effect transistor (MOSFET) sensor. This sensor type is sensitive to a number of organic compounds and selectivity can be achieved by using different kinds of gate metals and operating temperatures (Lundström et al 1990).

We have used an array of four different SnO₂ semiconductors and 10 MOSFET sensors, together with an infrared detector for CO₂ monitoring and ANN for data processing. In a previous evaluation of our equipment using subsamples from larger batches (Jonsson et al, *in press*), good predictions were achieved when classifying grain odor in oats and predicting the ergosterol content of wheat.

The main objective of the present investigation was to evaluate the accuracy of the apparatus at classifying off-odors, using true unknown test samples of grain that had previously been classified by grain inspectors. The overall goal was to further the development of an operational system for use in granaries.

MATERIALS AND METHODS

The electronic nose consists of three parts: an automatic sampling apparatus, a detector unit containing the sensors, and the ANN software.

¹Swedish Farmers Supply & Marketing Association, Box 30192, S-104 25 Stockholm, Sweden.

²Laboratory of Applied Physics, Linköping University, S-581 83 Linköping, Sweden.

³Department of Microbiology, Swedish University of Agricultural Sciences, S-750 07 Uppsala, Sweden.

⁴Corresponding author. Fax: 46 510 664 38.

Automatic Sampling Apparatus and Sampling Procedure

The apparatus (Fig. 1) was built by the Central Workshops, Swedish University of Agricultural Sciences, Uppsala, Sweden. It consists of a carousel containing 30 steel tubes for loading the grain samples (20 mm i.d., 300 mm long), and a heating unit (steel tube, 20 mm i.d., 400 mm long) provided with pneumatic valves at the top and bottom. During operation, grain samples in the sample tubes were transferred to the heating unit by pushing a Teflon plate with a hole 20 mm i.d. under the sample tube by means of a hydraulic piston. The sample then fell down to the heating unit below. The grain was heated in the heating unit to 65°C for 5 min. Sampling was done by pumping headspace air from the heating unit through the sensor chambers for 2 min.

The steel tube (1.5 mm i.d.) connecting the heating unit and the detector unit was heated to ≈65°C using a Teflon-coated heating wire (Habia Technoflor, Söderfors, Sweden) to prevent volatile compounds from condensing on the walls of the tube.

After sampling, the bottom valve was opened causing the sample to fall down to a collecting container underneath. The heating unit was cleaned with compressed air, and the connection to the detector unit was rinsed by pumping air through it (100 ml/min) for 10 min.

Detection Unit

All MOSFET sensors and the regulating circuits were produced at the Laboratory of Applied Physics, Linköping University, Sweden. They are *n*-channel field effect transistors with gates of thin, catalytically active metal films. The principle of operation can be summarized as follows. Hydrogen-containing compounds dissociate on the metal surface, and hydrogen diffuses through the metal film to form a dipole layer at the interface between metal and insulator. The dipole layer causes a voltage shift between the

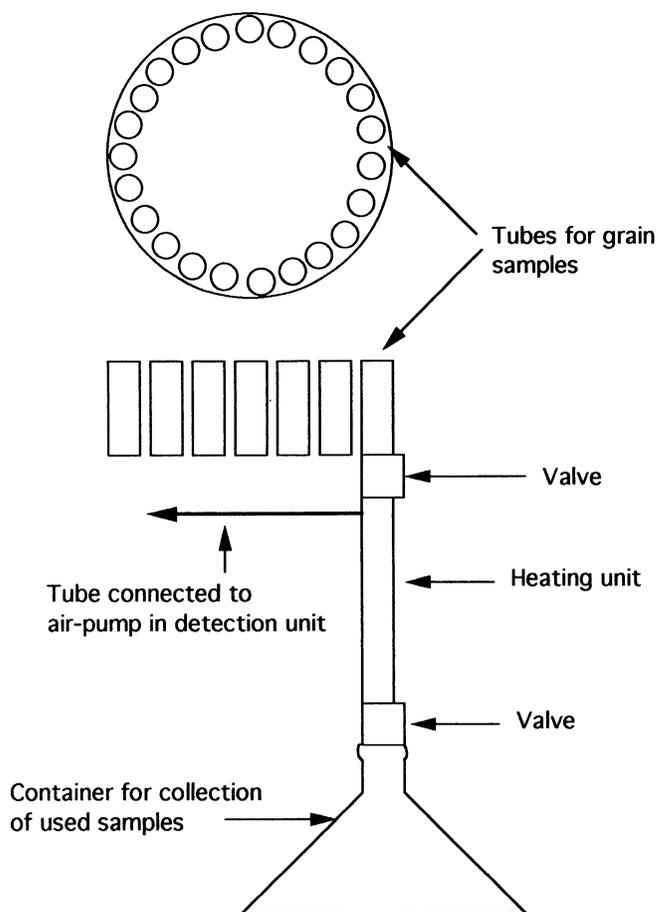


Fig. 1. Schematic display of sampling apparatus.

metal and the semiconductor. The external voltage on the metal will have to be lowered to maintain the original current through the transistor (Lundström et al 1993, Winquist et al 1993). The magnitude of the voltage decrease corresponds to the response of the sensor. The temperature, type, and thickness of the metal films influence the sensitivity of the devices toward different compounds. For thin, discontinuous metal gates, dipoles (or charges) on the metal or the insulator also give rise to a voltage shift (Lundström et al 1990). Two sets of five MOSFET sensors were constructed for this investigation, and each set was mounted in a separate chamber. The sensors were provided with metal films: palladium (6.5 and 35.0 nm), iridium (9.0 nm), platinum (5.0 nm), and a combination of platinum and palladium. One of the chambers was kept at 140°C during operation, and the other was kept at 175°C. Three of the MOSFET sensors, iridium (9.0 nm) 140°C, palladium (6.5 nm) 175°C, and palladium (35.0 nm) 175°C, showed reduced stability or weak responses when compared with the other sensors. Consequently, their signals were not used for data processing.

Four different SnO₂ based devices, constructed by Figaro Engineering Inc., Japan, were used: Taguchi sensors TGS 800, TGS 813, TGS 825, and TGS 880. They were mounted in a separate chamber and kept at 400°C during operation. The change in the voltage drop across a series resistor following exposure to a combustible gas was used as the sensor signal (Winquist et al 1993).

A Gascard CO₂ monitor (Edinburgh Sensors, Edinburgh, Scotland) was used for CO₂ measurements. A membrane pump (Bergman & Beving AB, Sweden) with an air flow of 100 ml/min was used to pump air through the sensor chambers during operation.

Grain Samples

During 1994, 235 samples of 1993 wheat (78), barley (59), and oats (98) were received from different granaries in Sweden. The samples had been graded by at least two inspectors at different locations. Most oat samples (91) were graded by three inspectors. Odors were described as either normal, musty, moldy, acid, sour, burnt, or foreign, and the intensities of off-odors were given as weak, pronounced, or strong (Statute Book 1991).

Each sample (50 ml) was run in the electronic nose at least once; some samples were run two or three times. A total of 386 response patterns were analyzed: 80 from wheat, 90 from barley, and 216 from oats. Three replicates were run on 39 of the oat samples; these patterns were data-processed individually as well as together with the others. Two of these replicates were run within a month of each other, and the third one was run about a month later. Most of these samples had either a good (normal) odor or a pronounced musty or moldy off-odor (Table I).

The reference grain was wheat of good microbial status and low moisture content (≈10%). Reference grain was run every sixth sample to check for changes in sensor responses during an individual run of the sampling apparatus. By using linear interpolation, responses in the sample runs were corrected based on the

TABLE I
Distribution of Odor Classifications for Two Inspectors
Classifying the Same 235 Grain Samples^a

Inspector A	Inspector B							Σ A
	Normal	Acid	Sour	Moldy	Musty	Burnt	Foreign	
Normal	89	6	0	1	13	1	0	110
Acid	2	5	0	0	1	2	1	10
Sour	3	2	0	0	2	0	1	8
Moldy	13	2	0	21	20	0	0	57
Musty	10	4	0	12	6	3	0	12
Burnt	1	1	0	0	4	6	0	12
Foreign	2	0	0	0	0	1	0	3
Σ B	120	20	0	34	46	13	2	

^a Samples on which the inspectors agreed are in bold numbers.

values obtained in the reference runs. Classification performance with correction was compared with that obtained without correction. Grain water content was calculated based on the weight loss after 12 hr at 103°C. All grain samples, including the reference grain, were stored at 2°C before analysis.

Pattern Recognition Routines

Supervised ANN training using back propagation was employed (Rumelhart et al 1986). During training, a set of patterns along with a given target output, which could be a single variable or class affiliation, were presented to the ANN. Weights in each processing unit were corrected, starting from the output layer and ending at the first hidden layer, until the deviations between achieved output and the preset values were acceptably small (Fig. 2). Preset membership to one of two possible classes was achieved by assigning the samples the output values 1 and 0. Output values for the other class were then 0 and 1. When more classes were used, more zeros were included in the output set of values.

The performance of the ANN when analyzing unknown patterns was then evaluated by means of a test set. For example, with two-group classification, two possible output values might be 0.85 and 0.15. These values can be interpreted as representing the relative degrees of membership to the two classes. In this case, the pattern showed a stronger affiliation to the former class. Degrees of class membership were calculated in the same way when more than two classes were used, and the class receiving the highest value was regarded as the predicted class.

The performance of the ANN when using a single output variable was estimated by calculating the correlation coefficient between predicted and measured values.

When running the grain samples, training was performed until a mean square error of 0.1 between the preset and the actual output values was achieved or a maximum of 10,000 iterations. Two hidden layers, containing six and four processing units, respectively, were used. Responses from each type of grain were analyzed separately and divided into a training set, containing two-thirds of the patterns, and a test set, containing the other third. The division into training and test sets was made three times so that each pattern was included in the test set once. This provides a kind of cross-validation, common for pattern-recognition routines. In cases where the same sample was run several times, all patterns were kept in either the training or test set during a training-testing procedure. Neural Network Explorer (Neural Ware, Pittsburgh, PA) was used for the ANN calculations.

Based on the descriptors given by the inspectors, the data sets were classified in three ways. 1) Division into four categories: moldy/musty, acid/sour, burnt, and normal. The classification of samples into groups was based on "majority decision." In cases where there were only two inspectors and they disagreed, the classification made last was used. When comparing the inspectors' decisions (Table I) we found that it was difficult to discriminate between moldy and musty samples. Thus, moldy and musty samples were combined into one class. The off-odor foreign was

rarely encountered, and there was not a single sample for which the majority of judges identified this off-odor. 2) Division into good (normal odor) and bad (having any kind of off-odor) categories. Only patterns from samples where all the inspectors agreed on this classification were included: 72 wheat patterns, 76 barley patterns, and 153 oats patterns. Ninety 90 patterns (30 samples) of the triplicate oat samples that were also treated individually. 3) Allotting each sample a mean value of the intensity of moldy/musty off-odor given by the inspectors. This may be more appropriate than categorization as a way of describing the presence of off-odors and, because grains are continuously being degraded during microbial growth, samples tend to range widely in quality from very good to very bad. Furthermore, almost all off-odor samples were either identified as moldy or musty (Table I). Disagreements among inspectors concerning classifications of other off-odors were common.

Thirty-four of the triplicated oat samples, the ones that were given the odor characteristics of moldy, musty, or normal, were used in this investigation. A moldy/musty odor was given 3 points if it was strong, 2 points if it was pronounced, 1 point if it was weak; those with normal odor received 0 points. The sum was calculated and divided by the number of inspectors.

Mean values of two and three runs in the electronic nose, respectively, were calculated. Using these data and the ANN procedure stated above, the correlation between the mean odor description given by the inspectors and the predicted moldy/musty odor was calculated. The effects of using mean values and correction for the response of reference grain on the correlations between measured and predicted moldiness were studied.

For comparison and to visualize the relationships between different samples and sensors, PLS (Partial Least Squares) was used. PLS is a tool in multivariate data analysis that produces projections in a few dimensions of a data matrix with many variables and objects (Wold 1989). The projections extract the information from the whole data table in a few dimensions (components) that can be presented graphically. The dimensions extracted in PLS are chosen to optimize the ability to use them to predict a dependent variable from many independent variables.

It has been used successfully earlier to predict sensory quality by using instrumental data (Martens 1985). The theoretical background of PLS has been extensively elaborated by Höskuldsson (1988). In our investigation, the sensors represent the independent variables, the mean values of moldy/musty odor represent the dependent variable, and the sample runs represent the objects. The program used was SIMCA-P (UMETRI AB, Umeå, Sweden).

RESULTS

Classification by Inspectors

Out of 235 samples, 54% were placed in the same category by the two inspectors (Table I). A large proportion of the disagreements were between descriptors that resemble each other, i.e., acid/sour and moldy/musty. When these two pairs of descriptors were combined, the percentage of agreement increased to 69%. When the number of classes was reduced to two (good and bad),

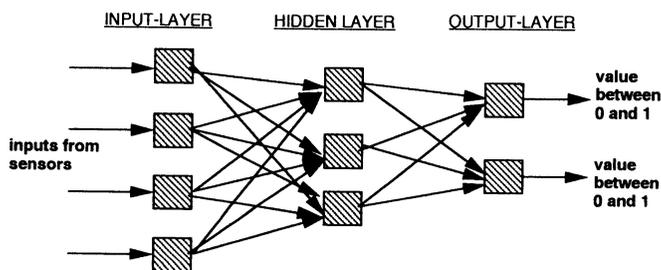


Fig. 2. Artificial Neural Network (ANN) with four input values, one hidden layer containing three nodes, and two output values.

TABLE II
Percentage Correct Classifications Obtained with the Artificial Neural Network (ANN) Classification System Using 2 or 4 Categories^a

Grain	Number of Categories			
	2	(n)	4	(n)
Wheat	89%	(80)	73%	(72)
Barley	88%	(90)	63%	(76)
Oats	91%	(216)	78%	(153)

^a System with 2 categories comprised samples classified as normal or with any off-odor. System with 4 categories comprised samples classified as normal, musty/moldy, acid/sour, or burnt.

the samples, approaching the 75% considered satisfactory by Dravnieks & Watson (1973). For ≈80% of the samples, the inspectors agreed on their classification as either good or bad. Thus, there seems to be little controversy as to what a good sample is, whereas it is more difficult to agree on how to describe the off-odor of a bad sample.

This problem was also encountered when using the electronic nose for classification. Classifications into four categories were correct ≈75% of the time. Most of the samples that had been characterized as burnt or acid/sour had been misclassified. The numbers of acid/sour and burnt samples may not have been sufficient to train the ANN properly.

Better results, with correct classifications ≈90% of the time, were achieved when grain samples were divided only into the categories good and bad. The highest percentage of correct classifications was achieved for oats, which was also the grain type with the highest number of response patterns, demonstrating that larger training sets increase the chances of covering the sample types that can occur in the test set. Support for lumping different odor descriptors together can be found in the literature. For example, high correlations were obtained between several different pairs of off-odor descriptors given by 10 trained panelists who described wheat samples stored at elevated moisture contents (Zawirska-Wojtasiak et al 1991).

High accuracy in predicting the mean moldy/musty off-odor was achieved in the present study. A correlation coefficient ($r = 0.89$) between perceived and predicted values for means of three runs was obtained. Predictions based on PLS were somewhat less accurate than those based on ANN.

In conclusion, the electronic nose can classify grain samples into good and bad with an accuracy of 90%, which is better than the level of agreement obtained between two inspectors classifying grains on different occasions. Bearing in mind that only those samples on which the inspectors had reached a consensus were used, one can conclude that the electronic nose and the human inspection are about equally effective in separating good and bad samples.

In view of the promising performance of the electronic nose and the efficient sampling apparatus tested, we are confident that an improved grain inspection routine, free from the health hazards and subjectivity involved in the human inspection routine, can be developed based on these technologies.

ACKNOWLEDGMENTS

We gratefully acknowledge Fredrik Winqvist for numerous valuable suggestions offered throughout this work, and Ingmar Lundström for critically reviewing the manuscript. Samples were provided and odors determined by several grain-receiving units in Sweden affiliated to Swedish Farmers Supply & Marketing Association. We would like to thank Layla Haglund, Margareta Hammarberg, Elisabeth Jansson, Inga Johansson, and Lotta B. Petersson, who work at these units. Odor determinations were also performed at the Swedish Cereal Laboratory, Svalöv, Sweden. Torsten Gumelius and Henrik Johansson are thanked for their assistance. Charlotte Essén is acknowledged for skillful assistance during the construction of the sampling apparatus, as is Bo Stenberg for assistance in performing PLS evaluations and for fruitful discussions concerning the method. The project has been financially supported by the Cerealia Foundation R&D, Järna, Sweden.

LITERATURE CITED

- AISHIMA, T. 1991. Aroma discrimination by pattern recognition analysis of responses from semiconductor gas sensor array. *J. Agric. Food Chem.* 39:752-756.
- BARTLETT, P. N., and GARDNER, J. W. 1992. Odour sensors for an electronic nose. Pages 31-51 in: *Sensors and Sensory Systems for an Electronic Nose*. J. W. Gardner and N. Bartlett, eds. Kluwer Academic Publishers; Dordrecht.
- BÖRJESSON, T., STÖLLMAN, U., ADAMEK, P., and KASPERSSON, A. 1989. Analysis of volatile compounds for detection of molds in stored cereals. *Cereal Chem.* 66:300-304.
- BÖRJESSON, T., STÖLLMAN, U., and SCHNÜRER, J. 1993. Off-odorous compounds produced by molds on oat meal agar: Identification and relation to other growth characteristics. *J. Agric. Food Chem.* 41:2104-2111.
- DRAVNIKS, A., and WATSON, C. A. 1973. Corn odor classification from low-resolution gas-chromatographic profiles of headspace volatiles. *J. Food Sci.* 38:1024-1027.
- HÖSKULDSSON, A. 1988. PLS regression methods. *J. Chemometrics* 2:211-228.
- JONSSON, A., WINQUIST, F., SCHNÜRER, J., SUNDGREN, H., and LUNDSTRÖM, I. In press. Electronic nose for microbial classification of grains. *Int. J. Food Microbiol.*
- LUNDSTRÖM, I., SPETZ, A., WINQUIST, F., ACKELID, U., and SUNDGREN, H. 1990. Catalytic metals and field-effect devices—A useful combination. *Sensors and Actuators B1:15-20*.
- LUNDSTRÖM, I., SVENSSON, C., SPETZ, A., SUNDGREN, H., and WINQUIST, F. 1993. From hydrogen sensors to olfactory images—Twenty years with catalytic field-effect devices. *Sensors and Actuators B13-14:16-23*.
- MAGA, J. A. 1978. Cereal volatiles, a review. *J. Agric. Food Chem.* 26:175-178.
- MARTENS, M. 1985. Sensory and chemical quality criteria for white cabbage studied by multivariate data analysis. *Lebensm. Wiss. Technol.* 18:100-104.
- RUMELHART, D. E., HINTON, G. E., and WILLIAMS, R. J. 1986. Learning internal representations by error propagation. Pages 318-362 in: *Parallel Distributed Processing*, Vol. 1. D. E. Rumelhart and J. L. McClelland, eds. MIT Press: Cambridge, MA.
- RYLANDER, R. 1986. Lung diseases caused by organic dusts in the farm environment. *Am. J. Ind. Med.* 10:221-227.
- SAMSON, R. A. 1985. Occurrence of moulds in modern living and working environments. *Eur. J. Epidemiol.* 1:54-61.
- STATUTE BOOK 1991. Statute Book of the Swedish Board of Agriculture. 1991:59, 7-8.
- STETTER, J. R. 1992. Chemical sensor arrays: Practical insights and examples. Pages 273-301 in: *Sensors and Sensory Systems for an Electronic Nose*. J. W. Gardner and N. Bartlett, eds. Kluwer Academic Publishers; Dordrecht.
- STETTER, J. R., FINDLAY, M. W., SCHROEDER, K. M., YUE, C., and PENROSE, W. R. 1993. Quality classification of grain using a sensor array and pattern recognition. *Anal. Chem. Acta* 284:1-11.
- WINQUIST, F., HÖRNSTEN, E., SUNDGREN, H., and LUNDSTRÖM, I. 1993. Performance of an electronic nose for quality estimation of ground meat. *Meas. Sci. Technol.* 4:1493-1500.
- WOLD, S. 1989. Multivariate data analysis: Converting chemical data tables to plots. *Intelligent Instruments and Computers Sept./Oct.:197-216*.
- ZAWIRSKA-WOJTASIAK, R., KAMINSKI, E., and ROGALSKA, M. 1991. Application of sensory profile analysis for quality evaluation of cereal grain during storage. Pages 351-366 in: *Proc. 3rd Wartburg Aroma Symp., Eisenach*. M. Rothe and H. P. Kruse, eds. Deutsches Institut für Ernährungsforschung: Potsdam-Renbrücke.

[Received July 17, 1995. Accepted April 16, 1996.]